Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **500,** 603–626 (2021)
Advance Access publication 2020 October 22

doi:10.1093/mnras/staa3269

Determining the systemic redshift of Lyman α emitters with neural networks and improving the measured large-scale clustering

Siddhartha Gurung-López,[1,2]★ Shun Saito,[1,3] Carlton M. Baugh,[4] Silvia Bonoli,[5] Cedric G. Lacey[4] and Álvaro A. Orsi[2,6]

[1]*Institute for Multi-messenger Astrophysics and Cosmology, Department of Physics, Missouri University of Science and Technology, 1315 N. Pine St., Rolla, MO 65409, USA*
[2]*Centro de Estudios de Física del Cosmos de Aragón, Plaza San Juan 1, piso 2, E-44001 Teruel, Spain*
[3]*Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, The University of Tokyo, Kashiwanoha, Kashiwa, Chiba 277-8583, Japan*
[4]*Institute for Computational Cosmology, Durham University, Durham, DH1 3LE, UK*
[5]*DIPC, Manuel Lardizabal Ibilbidea, 4, E-20018 San Sebastian, Spain*
[6]*PlantTech Research Institute Limited, South British House, 4th Floor, 35 Grey Street, Tauranga 3110, New Zealand*

Accepted 2020 October 11. Received 2020 September 6; in original form 2020 May 7

ABSTRACT

We explore how to mitigate the clustering distortions in Lyman α emitter (LAE) samples caused by the misidentification of the Lyman α (Ly α) wavelength in their Ly α line profiles. We use the Ly α line profiles from our previous LAE theoretical model that includes radiative transfer in the interstellar and intergalactic mediums. We introduce a novel approach to measure the systemic redshift of LAEs from their Ly α line using neural networks. In detail, we assume that for a fraction of the whole LAE population their systemic redshift is determined precisely through other spectral features. We then use this subset to train a neural network that predicts the Ly α wavelength given an Ly α line profile. We test two different training sets: (i) the LAEs are selected homogeneously and (ii) only the brightest LAE is selected. In comparison with previous approaches in the literature, our methodology improves significantly the accuracy in determining the Ly α wavelength. In fact, after applying our algorithm in ideal Ly α line profiles, we recover the clustering unperturbed down to 1 cMpc h^{-1}. Then, we test the performance of our methodology in realistic Ly α line profiles by downgrading their quality. The machine learning technique using the uniform sampling works well even if the Ly α line profile quality is decreased considerably. We conclude that LAE surveys such as HETDEX would benefit from determining with high accuracy the systemic redshift of a subpopulation and applying our methodology to estimate the systemic redshift of the rest of the galaxy sample.

Key words: radiative transfer – galaxies: high-redshift – intergalactic medium.

1 INTRODUCTION

Since the detection of the first galaxies emitting Lyman α radiation more than 20 yr ago (e.g. Steidel et al. 1996; Hu, Cowie & McMahon 1998; Rhoads et al. 2000; Malhotra & Rhoads 2002), Lyα radiation (with wavelength $\sim$1215.68 Å in rest frame) has been used as a successful tracer in the local (Henry et al. 2018; Orlitová et al. 2018) and of the high-redshift Universe (Ouchi et al. 2008; Steidel et al. 2010, 2011; Jones, Stark & Ellis 2012; Chonis et al. 2013; Erb et al. 2014; Trainor et al. 2015; Guaita et al. 2017; Matthee et al. 2017; Oyarzún et al. 2017; Caruana et al. 2018), detecting galaxies even at the epoch of reionization (Sobral et al. 2015; Ouchi et al. 2018; Shibuya et al. 2018). Ongoing cosmological galaxy surveys, such as the Hobby–Eberly Telescope Dark Energy Experiment (HETDEX; Hill et al. 2008; Adams et al. 2011) and the Javalambre Physics of the Accelerating Universe Astrophysical Survey (Benitez et al. 2014; Bonoli et al. 2020), aim at unveiling the nature of the dark energy using LAEs at the high-redshift Universe. One of the most useful tools to extract cosmological information from galaxy surveys is the galaxy clustering (e.g. Shoji, Jeong & Komatsu 2009). Therefore, understanding the spatial distribution of LAEs has become more important than ever before.

The complexity in understanding the LAE clustering resides in the radiative transfer of Ly α photons inside neutral hydrogen (Harrington 1973; Neufeld 1990). In first place, Ly α photons are emitted in the H II regions around OB-*type* stars. Then they have to cross the interstellar medium (ISM), then the circumgalactic medium (CGM), and the intergalactic medium (IGM) until they finally reach our observatories. In all these three mediums there is neutral hydrogen, and therefore they are optically thick to Ly α radiation. Inside galaxies, it is commonly thought that Ly α photons escape through outflow that modify the Ly α flux and line profile (e.g. Ahn, Lee & Lee 2000; Zheng & Miralda-Escudé 2002; Ahn 2003; Verhamme, Schaerer & Maselli 2006; Orsi, Lacey & Baugh 2012; Gronke et al. 2016; Gurung-López, Orsi & Bonoli 2018a). Then, in the CGM the Ly α radiation is spread around the galaxy creating the so-called Ly α haloes (e.g. Zheng et al. 2010; Behrens et al. 2017; Leclercq et al. 2017). Finally, the Ly α radiation enters inside the IGM, where, to a first approximation, the radiation bluewards

★ E-mail: gurung.lopez@gmail.com

the Ly α wavelength is absorbed (e.g. Laursen, Sommer-Larsen & Razoumov 2011; Zheng et al. 2011; Byrohl, Saito & Behrens 2019; Gurung-López et al. 2020).

Furthermore, the clustering property of LAEs can be sensitive to the selection function that is typically determined by the flux threshold. For example, Ly α radiation is very sensitive to dust. Therefore, galaxies with low metallicities are preferentially observed as LAEs (Sobral et al. 2018). This translate into a lower clustering amplitude (Gurung López et al. 2018a), as galaxies with lower metallicity exhibit are hosted in smaller dark matter haloes (e.g. Lacey et al. 2016). Moreover, it has been pointed out that the large-scale properties of the IGM might play a role on the selection function of LAEs (Zheng et al. 2010, 2011). This could distort the clustering of LAEs and reduce the accuracy of LAE cosmological surveys (Wyithe & Dijkstra 2011). However, there is still debate in the community whether if there is a large-scale IGM coupling with the observed Ly α luminosity (Gurung-López et al. 2020) or not (Behrens et al. 2017). These facts contribute to the complexity of understanding the radiative transfer of the Ly α radiation and its impact on the clustering statistics.

A key challenge lies in determining the systemic redshift of an LAE only from their Ly α line profile. Due to the radiative transfer, the peak of the Ly α line rarely matches the Ly α wavelength. In Verhamme et al. (2018), authors probed that taking into account the radiative transfer in the ISM was crucial to improve the determination of the systemic redshift of LAEs. In their work, they used a relation between the shift of the line peak from Ly α and the width of the line to correct for the redshift due to the radiative transfer into the ISM. However, this recipe assumed the commonly used Thin Shell toy model for the outflow geometry. In fact, the relation between the width of the line and peak shift depends strongly on the outflow geometry used (Gurung-Lopez et al. 2018a). More recently, Muzahid et al. (2019) linked the LAEs star formation rate (SFR) to the line peak offset, also, getting better results than when assuming that the peak of the Ly α line is the Ly α wavelength.

Along a similar line, Byrohl et al. (2019) studied how the incorrect systemic redshift determination distorts the clustering of LAEs along the line of sight. In detail, an inaccurate Ly α wavelength determination is translated into an imprecise redshift, thus into an uncertainty in its radial position in redshift space. This causes that LAEs look like they are more spread along the line of sight than what they actually are. In fact, this shuffling in the LAE radial position can be interpreted as a extra random radial velocity dispersion component, which translates into an additional 'Finger-of-God' suppressing the apparent clustering along the line of sight. Furthermore, Byrohl et al. (2019) found that the clustering distortion was mitigated after correcting the shift of the peak with different recipes, such as the those in Verhamme et al. (2018). However, the developed recipes in the literature (Verhamme et al. 2018; Byrohl et al. 2019; Muzahid et al. 2019) to estimate the Ly α wavelength have limitations as the dispersion of the estimated Ly α wavelength was around 1 Å (in rest frame). Such a large scatter introduces significant distortions to the apparent clustering of LAEs on scales ~ 5 cMpc h^{-1} in the monopole and up to $k \gtrsim 0.1\,h$ cMpc^{-1} in Fourier space, as we will explicitly show in this paper.

We propose a novel approach to determine the systemic redshift from the Ly α line profiles of LAEs using neural networks. This is motivated by the fact that there must be information on the Ly α wavelength in an entire spectral range of Ly α line profiles. We explore whether or not a given survey that observes LAEs only through Ly α emission (e.g. HETDEX; Hill et al. 2008) could benefit from acquiring a subsample with a systemic redshift (without using Ly α), for example, by H α observations. Then, this subset could be used to train a neural network to predict the Ly α wavelength in the rest of the main LAE population. In this work, we train different neural networks using subsets of the Ly α line profiles computed by our model to predict their Ly α wavelength.

This work is part of a series of papers studying the impact of the Ly α radiative transfer on the observed properties of LAEs. In our first work (Gurung-López et al. 2019a), we focused on the Ly α radiative transfer (RT) taking place in the ISM. There, we found that LAEs are a very peculiar population that exhibits a tight balance between SFR and metallicity. Then, in our second work (Gurung-López et al. 2020) we implemented the Ly α in the IGM and we focused on the different selection effects introduced by it. There, we studied the impact in the clustering on large scales due to the IGM-LAE coupling. In this third work, we analyse the properties of the Ly α stacked line profiles and study the impact on small scales of the misidentification of the Ly α wavelength. None the less, we emphasize that we adopt this simulation set for a proof of concept, and that our approach can be generally applied to any LAE spectroscopic observation, in principle.

We briefly describe our model in Section 2, while for a more extensive description we refer the reader to Gurung-López et al. (2019a, 2020). Then, in Section 3 we study the properties of the Lyα stacked line profiles. In Section 4, we describe the different methodologies used in this work to identify the systemic redshift of LAEs from their observed Ly α line profiles. Then, in Section 5 we describe the effects of the Ly α misidentification in ideal line profiles. Meanwhile, in Section 6 we artificially reduce the quality of our Ly α line profiles and analyse the effects of the Ly α misidentification in realistic Ly α line profiles. Then, we discuss our result in Section 7. Finally, we make our conclusions in Section 8.

Throughout this paper, all the properties related to Ly α line profiles are given in length units in the rest frame of the LAEs.

2 LAE THEORETICAL MODEL

In this work, we adopt the LAEs simulated with a semi-analytical model in Gurung-López et al. (2019a, 2020). In this section, let us briefly describe the LAE model but emphasize on modelling the spectrum around the Ly α emission. Our LAE model is based on four main ingredients:

(1) The dark matter N-body simulation, P-MILLENNIUM (Baugh et al. 2019), that imprints the hierarchical growth of structures in the ΛCDM scenario. This state-of-the-art cosmological simulation consists in 5040^3 dark matter particles with mass of $1.061 \times 10^8\,\mathrm{M}_\odot\,h^{-1}$ distributed in a volume of $L^3_{\rm box} = (542.16\,\mathrm{cMpc}\,h^{-1})^3$. P-MILLENNIUM uses cosmological parameters: $H_0 = 67.77\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\Lambda = 0.693$, $\Omega_\mathrm{m} = 0.307$, $\sigma_8 = 0.8288$, consistent with Planck Collaboration XIII (2016).

(2) The model of galaxy formation and evolution, GALFORM (Cole et al. 2000; Lacey et al. 2016; Baugh et al. 2019). In short, GALFORM populates galaxies and gases within the dark matter haloes and tracks their evolution through the cosmic history. GALFORM follows recipes to estimate a whole bunch of galaxy properties such as metallicity or SFR. These recipes are calibrated to fit several observables, such as, the optical and near-infrared luminosity functions at $z = 0$ and its evolution up to $z = 3$, the H I mass function at $z = 0$, the submm galaxy number counts and their redshift distributions among others. Galaxies in GALFORM exhibit two components: the discs, where the quiescent star formation takes place and the bulges, where the strong star formation bursts take place. Each of these pieces exhibit different properties (such as metallicity, etc.).

(3) The Ly α radiative transfer in the ISM is implemented through the PYTHON open source code, FLAREON (Gurung-López, Orsi & Bonoli 2019b). FLAREON is based on a pre-computed grid of outflow models using LYART (Orsi et al. 2012), spawning a wide range in neutral hydrogen column density ($N_{\rm H}$), outflow expansion velocity ($V_{\rm exp}$), and dust optical depth ($\tau_{\rm a}$). By using different machine learning and multidimensional interpolation algorithms, FLAREON predicts the Ly α escape fraction $f_{\rm esc}$ and line profile $\phi(\lambda)$ with high accuracy for different outflow geometries. Here, we will focus on the 'Thin Shell' and 'Galactic Wind' outflow geometries.

Throughout this work, we define the escape fraction of a given medium as the ratio between the flux injected into the medium and that emerges from it. For example, for the ISM escape fraction, $f_{\rm esc}^{\rm ISM} = L_{{\rm Ly}\,\alpha,\,{\rm ISM}}/L_{{\rm Ly}\,\alpha,\,0}$ where $L_{{\rm Ly}\,\alpha,\,0}$ and $L_{{\rm Ly}\,\alpha,\,{\rm ISM}}$ are the intrinsic Ly α luminosity and the luminosity after passing through the ISM region, respectively. Also, in our convention the line profile, $\phi(\lambda)$, is normalized as $\int \phi(\lambda){\rm d}\lambda = 1$.

In practice, our model links the galaxy properties predicted by GALFORM to outflow features through simple recipes (Gurung-López et al. 2019a). In this way, each component (disc and bulge) in each galaxy has a different parameter set of $\{N_{\rm H}, V_{\rm exp}, \tau_{\rm a}\}$ through their SFR, metallicity, cold gas mass, and stellar mass. Then, we use FLAREON to compute an Ly α line profile and an escape fraction for each galaxy and component.

(4) The radiative transfer in the IGM is implemented by computing the optical depth of Ly α photons in the line of sight (fixed along Z-axis) between the observer and each galaxy. Therefore, the IGM transmission depends on the particular properties of the environment of each galaxy, and in particular, on the IGM density ρ, its density gradient along the line of sight $\partial_Z \rho$, the IGM line-of-sight velocity V_Z, and its gradient $\partial_Z V_Z$.

Our model provides the Ly α line profile and luminosity is computed by convolving the Ly α line profile emerging from the ISM with the IGM transmission curve for each component of a galaxy. In practice, the line profile and luminosity are given by

$$\phi(\lambda) = \frac{L_{{\rm Ly}\,\alpha}^{\rm Disc}\phi_{\rm ISM}^{\rm Disc} + L_{{\rm Ly}\,\alpha}^{\rm Bulge}\phi_{\rm ISM}^{\rm Bulge}}{L_{{\rm Ly}\,\alpha}}, \quad (1)$$

and

$$L_{{\rm Ly}\,\alpha} = L_{{\rm Ly}\,\alpha}^{\rm Disc} + L_{{\rm Ly}\,\alpha}^{\rm Bulge}, \quad (2)$$

where the luminosity for each component is evaluated as

$$L_{{\rm Ly}\,\alpha}^{X} = L_{{\rm Ly}\,\alpha,0}^{X} f_{\rm esc}^{\rm ISM,X} f_{\rm esc}^{\rm IGM,X}, \quad (3)$$

with $X = \{{\rm Disc}, {\rm Bulge}\}$. $f_{\rm esc}^{\rm IGM}$ is the escape fraction from the IGM.

Furthermore, LAEs are defined as galaxies with an Ly α emission line exhibiting a high contrast to the galaxy continuum. It is usually found in the literature that for a galaxy to be considered an LAE, it must exhibit a rest-frame equivalent width ${\rm EW}_0 > 20$ Å (e.g. Gronwall et al. 2007; Konno et al. 2018). In this work, we follow this criteria. In particular, we compute the ${\rm EW}_0$ for each galaxy as a function of its continuum luminosity per unit of wavelength around Ly α wavelength $L_{\rm c}$ as

$${\rm EW}_0 = L_{{\rm Ly}\,\alpha}/L_{\rm c}, \quad (4)$$

where $L_{\rm c}$ is directly provided by GALFORM and it is based on the full evolution of the stellar population given a galaxy.

The free parameters in the model, which depend on the outflow geometry model, are adjusted so that the simulated LAEs reproduce the observed Ly α luminosity function at their corresponding redshift, as described in Gurung-López et al. (2020).

Finally, the samples studied here are selected by making a number density cut of $4 \times 10^{-3} (h\,{\rm cMpc})^{-3}$ in Ly α luminosity. Given our simulation volume, each LAE population (combination of redshift and outflow geometry) is compound by 637 444 galaxies.

3 STACKED LYMAN α LINE PROFILES

Since the main goal of this paper is to study the determination of the systemic redshift of LAEs from their Ly α line profile, we present here the detailed properties of the stacked Ly α line profile. First, we focus on how Ly α radiative transfer impacts the observed Ly α stacked line profile in our model. Then, we discuss the Ly α line profiles as a function of different galaxy and IGM properties to understand how they influence the stacked line profile.

Throughout this work, we compute the stacked Ly α line profiles in a consistent manner. First, we normalize all the Ly α line profiles so that they have an area of unity, i.e. compute $\phi(\lambda)$. Then, we evaluate the stacked line profile, $\langle\phi(\lambda)\rangle$, as the median of the line profile collection.

3.1 The impact of radiative transfer

In Fig. 1, we compare the stacked Ly α profile $\langle\phi(\lambda)\rangle$ before (dashed thin) and after (solid thick) being processed by the IGM. In general, at low redshift ($z = 2.2$, 3.0) the IGM tends to absorb blue photons, i.e. $\lambda < \lambda_{{\rm Ly}\,\alpha}$, while it does not affect less for redder photons. Meanwhile, at $z = 5.7$ the IGM optical depth to Ly α is much greater and it affects to the stacked line profile up to wavelength 2 Å redder than Ly α.

Furthermore, the impact of the IGM in the stacked line profiles is different for each geometry. We find that at $z = 2.2$ and $z = 3.0$, the IGM affects more the Thin Shell geometry than the Galactic Wind. In particular, the blue peak that is present before the line profiles is processed by the IGM (especially at $z = 3.0$), is mostly vanished after the RT in the IGM. Meanwhile, at $z = 5.7$ the IGM affects similarly both cases. This is consistent with the values of the Ly α IGM escape fraction in our model (see figs 10 and 11 in Gurung-López et al. 2020). These findings originate from the differences in the family of line profiles generated by each outflow geometry, as the IGM absorption depends on the wavelength of the photons, as discussed in detail in Gurung-López et al. (2020).

Our model predicts different shape of $\langle\phi(\lambda)\rangle$ for the Thin Shell and Galactic Wind models at all the redshift bins studied in this work. In general, we find that the stacked line profile for the Thin Shell model is bluer than that for Galactic Wind. Meanwhile, at each different epoch the stacked line profiles for two outflow geometry models differ in unique fashions. For example, at $z = 2.2$ the peaks of $\langle\phi\rangle$ of both geometries match, but the Thin Shell exhibits a broader $\langle\phi\rangle$ (∼3 Å) than the Galactic Wind(∼1.5 Å). At $z = 3.0$, the width of the stacked profiles are comparable between both geometries, but the Galactic Wind profile is more redshifted than the Thin Shell one. Another difference between our two outflow models is that at redshift of 2.2 and 3.0, the Thin Shell exhibits a weak peak bluer than Ly α, while the Galactic Wind lacks this blue peak. Finally, at $z = 5.7$ the Thin Shell profile is broader than the Galactic Wind one.

The differences between the 'Thin Shell' and 'Galactic Wind' outflow geometries arise due to two facts. First, the distribution of the ISM parameters, $\{V_{\rm exp}, \tau_a, N_{\rm H}\}$, of the LAEs is different for both geometries (see figs A1 and A2 of Gurung-López et al. 2020). Second, the radiative transfer in each geometry leads to a different line profile, even for the same parameter set of $\{V_{\rm exp}, \tau_a, N_{\rm H}\}$. For example, the 'Thin Shell' geometry model is more prone to exhibit a blue peak than the 'Galactic Wind' model (Gurung-López et al.

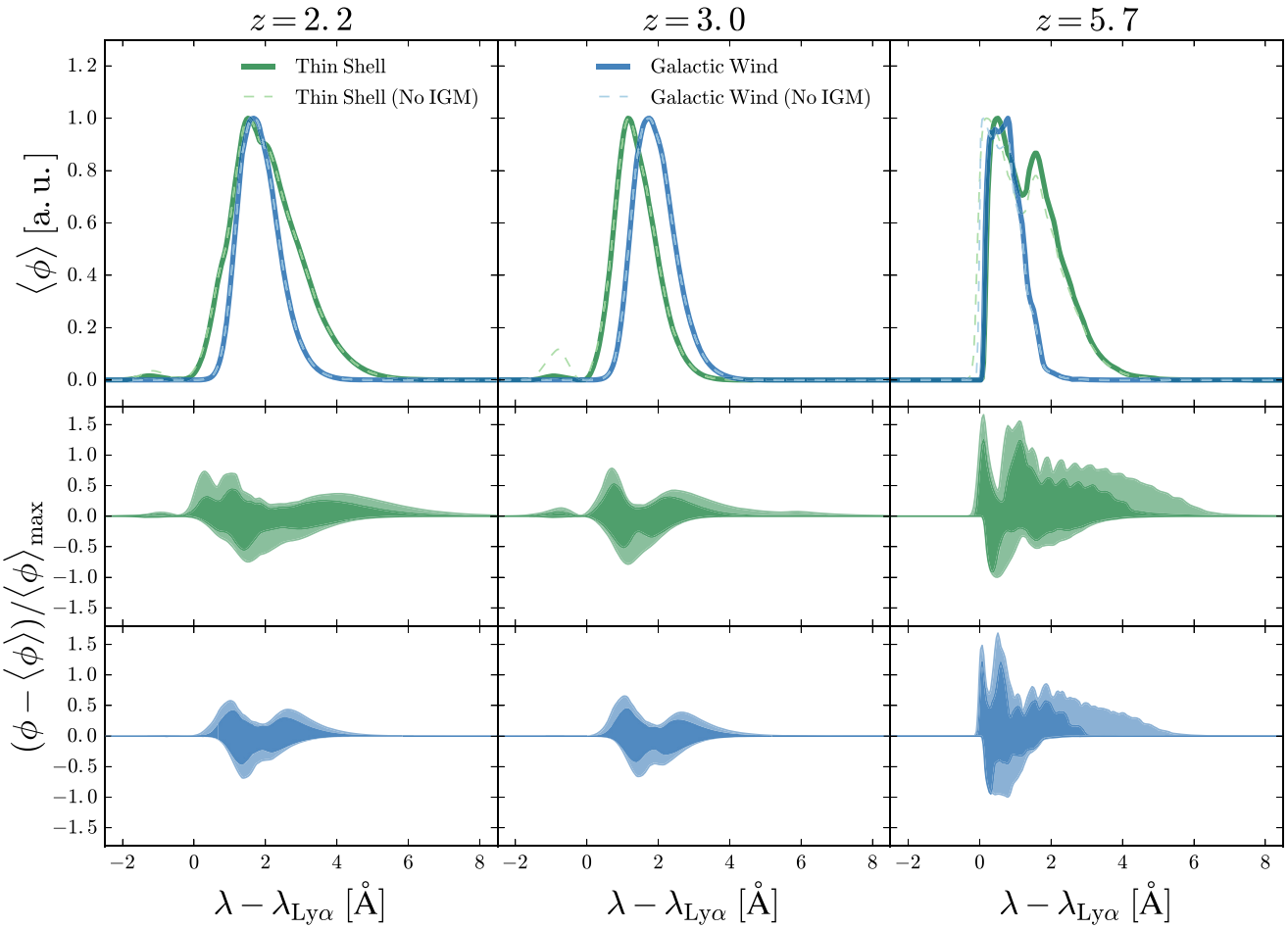

Figure 1. Top: Stacked Ly α line profile at redshift 2.2, 3.0, 5.7 from left to right. The Thin Shell geometry is displayed in thick green lines and the Galactic Wind in blue. Our default models, including the IGM Ly α absorption, are shown in thick lines, while models without the IGM implementation are shown in thin dashed lines. The stacked line profiles are normalized so that their global maximums match unity. Middle: Difference between individual and stack line profiles relative to the maximum of the stack line profile for the Thin Shell geometry including the IGM. The inner dark (light) regions show the area between the percentiles 16th and 84th (5th and 95th). Bottom: Same as middle but for the Galactic Wind outflow geometry including the IGM.

2019b), as we see in the stacked profile at redshifts 2.2 and, in particular, at $z = 3.0$.

In summary, the radiative transfer impacts the shape of the stacked profile including the peak position and the width from the ISM to IGM scales in a non-trivial manner.

In the lower panels of Fig. 1, we show the diversity of line profiles for our models including the IGM. We display the relative difference to the maximum of the stack line profile of the line profile population. In the middle row, we display the Thin Shell while the Galactic Wind is shown in the bottom. The dark (light) shaded region shows the area between the percentile 16th and 84th (5th and 95th) of the line profile distribution. We find that the variety of lines is similar at redshifts 2.2 and 3.0. At these redshifts, the Thin Shell geometry has some dispersion at wavelengths lower then Ly α. This indicates that in these models, some lines exhibit a blue peak even after the IGM absorption. Meanwhile, the models using the Galactic Wind lack emission at bluer frequencies than Ly α. At redshift 5.7, the diversity of lines increases, specially for the Thin Shell. The morphology of the relative difference (shaded areas) changes with respect to $z = 2.2$ and 3.0, getting wider in wavelength and exhibiting fluctuations with larger amplitudes.

3.2 Dependence of the stacked profile on galaxy and environmental properties

So far we have considered the stacked profile for all LAEs in our simulation. In this section, we split the Ly α stacked line profile according to different galaxy and IGM properties. To this end, for a given galaxy or IGM property X, we compute the percentiles 33.33 ($Q(33)$) and 66.66 ($Q(66)$). Then, we split the LAEs into three subsamples of the same size containing the galaxies with the lowest ($Q(0) < X < Q(33)$), intermediate ($Q(33) < X < Q(66)$), and the greatest ($Q(66) < X < Q(100)$) values of X.

3.2.1 Imprints of the galaxy properties

In Fig. 2, we split the Ly α stacked line profile as a function of the SFR, rest-frame Ly α equivalent width, Ly α luminosity, and stellar mass (from left to right). The Ly α stack line profile of each subpopulation is normalized to the maximum of the stacked line profile of the complete LAE sample, $\langle\phi\rangle_{\rm max}$. Here, we show the snapshot at redshift 3.0 only, but have confirmed that the other two redshift bins ($z = 2.2$ and 5.7) exhibit similar trends.

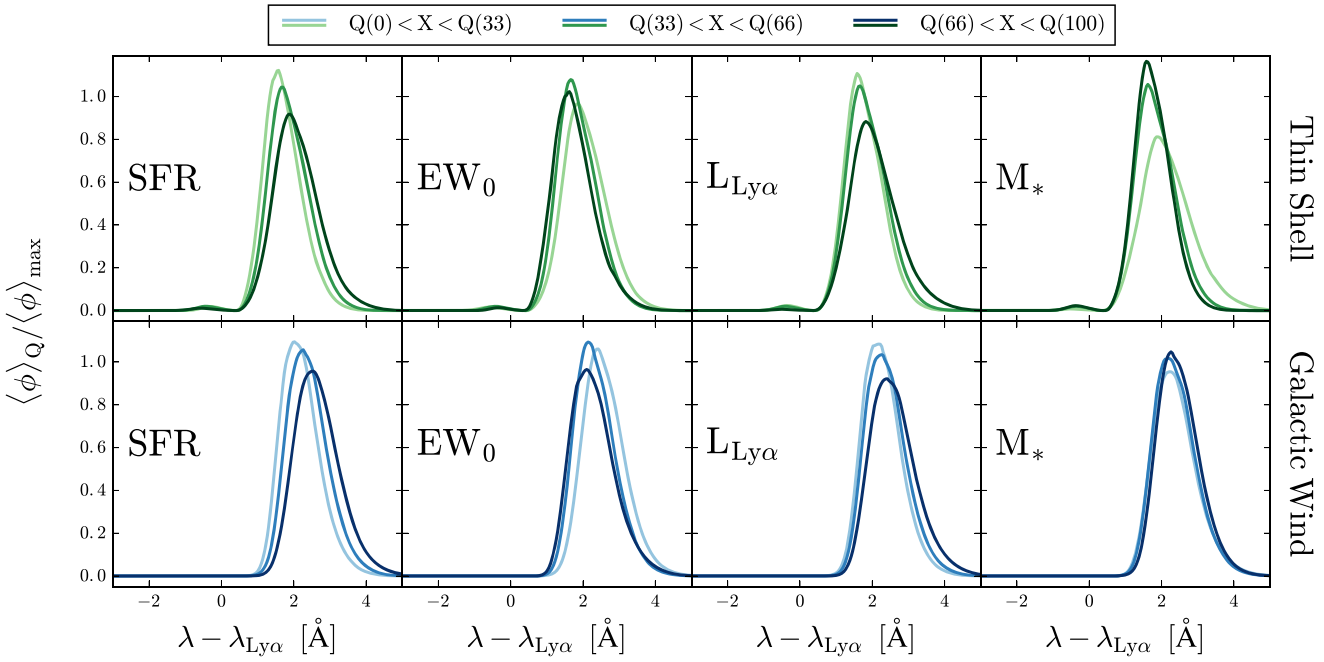

Figure 2. Break down of the Ly α stacked line profiles at $z = 3.0$ as a function of total SFR, Ly α rest-frame equivalent width, Ly α luminosity, and stellar mass from left to right. Top panels show the model with the Thin Shell outflow geometry, while bottom panels show the model using the Galactic Wind geometry. We rank our main LAE population by each of these galaxy properties and split it by the percentiles 33 ($Q(33)$) and 66 ($Q(66)$). The stacked line profiles are colour coded according to the samples they are displaying. The lighter lines show the samples with the lowest values of the galaxy property X ($Q(0) < X < Q(33)$). The darkest lines show the samples with the highest values of each galaxy property ($Q(66) < X < Q(100)$). Meanwhile, the intermediate values ($Q(33) < X < Q(66)$) are shown in intermediate colours. The Ly α stack profile of each subpopulation, $\langle\phi\rangle_Q$, is normalized to the maximum of the stacked line profile of the complete LAE population, $\langle\phi\rangle_{max}$.

Overall, our model predicts that the stacked Ly α line profile properties depend on the galaxy properties. For instance, the peak of the Ly α line profile correlates positively with the SFR and $L_{Ly\alpha}$ for both outflow geometries. The peak of the line profile anticorrelates with EW_0 for both outflow geometries. Interestingly, the dependence of the stacked Ly α line profile on the stellar mass behaves differently for two outflow geometries. In Thin Shell, the peak anticorrelates with M_*, while there is no apparent trend in Galactic Wind.

We also find that not only the peak position, but also the shape of the stacked line profile changes through the dynamical range of these galaxy properties. In particular, we find that for both outflow geometries, the stacked line profile becomes broader at higher SFR and $L_{Ly\alpha}$ values, while it shrinks for high values of EW_0. Meanwhile, when the 'Thin Shell' is implemented, increasing the stellar mass leads to broader stacked profiles. In contrast, the width of the stacked line profile using the 'Galactic Wind' remains constant through the stellar mass dynamical range.

These non-trivial differences between Thin Shell and Galactic Wind are consequences of the complicated interplay between the properties of galaxies and outflows (see equations XX and XX in Gurung-López et al. 2020). This highlights the importance of having different outflow models to model the RT in the ISM.

3.2.2 *Imprints of the IGM*

Here, we study how the different large-scale IGM properties change the observed Ly α line profile. In order to do so, we split the Ly α stacked line profile $\langle\phi\rangle$ by IGM properties for both outflow geometries. In Fig. 3, we show the difference between the stacked line profile of the full LAE population and the split samples. The IGM properties used for dividing the LAE population are density ρ, density gradient along the line of light $\partial_Z\rho$, velocity along the line of sight V_Z, and its gradient along the line of sight $\partial_Z V_Z$ from left to right. These properties were computed in a regular grid of cubic cells of 2 cMpc h^{-1} side, as in Gurung-López et al. (2020). Here, we focus on the snapshot at $z = 5.7$, where IGM is optically thicker to Ly α photons than at $z = 2.2$ and $z = 3.0$. Note that we also find the same trends at $z = 2.2$ and 3.0, but with a lower amplitude.

We find that, although differences are tiny ($\lesssim 10$ per cent compared to $\langle\phi\rangle_{max}$), both outflow geometries exhibit the same trends with a clear dependence of the IGM properties. This suggests that there is a smooth dependence between the IGM properties and the stacked line profile. For example, the higher the IGM density, the more flux is absorbed at bluer wavelengths. This causes that the observed stacked line profile is slightly more redshifted in high IGM density regions. Also, our model predicts that LAEs located in regions with high $\partial_Z\rho$, V_Z, and $\partial_Z V_Z$ exhibit a bluer stacked line profile, while the opposite is true for low values of these IGM properties.

The trends in the stacked line profile are in agreement with the our previous work (Gurung-López et al. 2020), where the IGM transmission generally anticorrelates with ρ and positively correlates with $\partial_Z\rho$, V_Z, and $\partial_Z V_Z$ (Gurung-López et al. 2020). We also showed that the bluer the wavelength around Ly α, the more sensitive the line profile is to IGM absorption (see fig. 5 in Gurung-López et al. 2020). Combining these two facts, the LAEs lying in regions with lower IGM transmission will exhibit a redder line profile than LAEs lying in regions with higher transmission, which is consistent with Fig. 3.

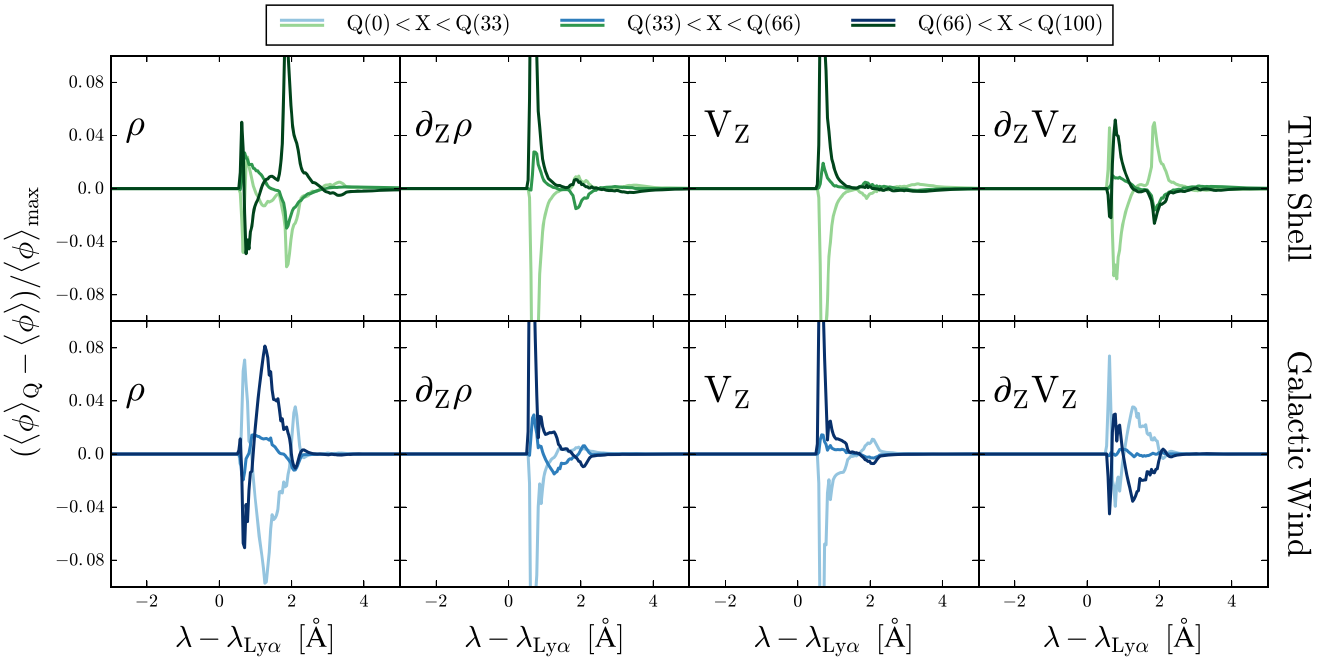

Figure 3. Break down of the Ly α stacked line profiles at redshift $z = 5.7$ as a function of the IGM large-scale properties (density, density gradient along the line of sight, velocity along the line of sight, and its gradient along the line of sight from left to right). For each of these IGM properties, we rank and divide the LAE population in three samples of the same size. Here, we show the difference between the stack of these subsamples and the stacked line profile of the full LAE sample. In the top (bottom) panels, we show the Thin Shell (Galactic Wind). The colour code is the same as in Fig. 2.

4 DETERMINING THE REDSHIFT OF LAES

As we showed in the previous section, the wavelength of photons initially emitted at the Ly α wavelength changes as they travel through the ISM and the IGM. In this way, the Ly α line profiles are modified in a non-trivial way (e.g. Zheng et al. 2011) by the Ly α RT. This complicates the determination of the Ly α wavelength from an observed Ly α line profile (Verhamme et al. 2018; Byrohl et al. 2019). In general, in each Ly α line profile, the true Ly α wavelength ($\lambda_{\mathrm{Ly}\alpha}$) and the wavelength set as Ly α ($\lambda^{\mathrm{Obs}}_{\mathrm{Ly}\alpha}$) can differ, as we show below.

In the following, we introduce the different methods that we use through this work to find the Ly α wavelength ($\lambda^{\mathrm{Obs}}_{\mathrm{Ly}\alpha}$) directly from the Ly α line profile. First, in Section 4.1 we describe two different methods to retrieve $\lambda^{\mathrm{Obs}}_{\mathrm{Ly}\alpha}$ that have been already used in the literature. These algorithms depend on line profile characteristics such as the width and the position of the global maximum of the Ly α line. Then, in Section 4.2 we introduce a novel method that makes use of the full Ly α line profile in order to predict $\lambda^{\mathrm{Obs}}_{\mathrm{Ly}\alpha}$ through neural networks.

4.1 Standard methodologies

(i) **GM** (global maximum): This is the simplest method to assign an Ly α wavelength. Basically, the position of the global maximum ($\lambda_{\mathrm{Ly}\alpha,\,\mathrm{Max}}$) is set as the Ly α wavelength, i.e.

$$\lambda^{\mathrm{Obs}}_{\mathrm{Ly}\alpha} = \lambda_{\mathrm{Ly}\alpha,\mathrm{Max}}. \tag{5}$$

(ii) **IC** (intensity centre) : This method assigns the centroid of the line as the Ly α wavelength, i.e.

$$\lambda^{\mathrm{Obs}}_{\mathrm{Ly}\alpha} = \frac{\sum \phi(\lambda)\lambda}{\sum \phi(\lambda)}. \tag{6}$$

Several works in the literature (e.g. Steidel et al. 2010; Rudie, Steidel & Pettini 2012) use a similar approach to IC to estimate the redshift of LAEs.

(iii) **GM-F**: This method takes into account that the Ly α photons tend to be redshifted as they escape the galaxies through outflows. As a result, the position of the red peak is shifted from the Ly α frequency. This shift depends on the outflow properties and can be related to the full width at half-maximum (FWHM) of the red peak (Verhamme et al. 2018; Gurung-López et al. 2019b). As we have shown (see Section 3), the Ly α line profiles predicted by our model are clearly dominated by a prominent red peak and only a faint blue peak is found. Therefore, $\lambda_{\mathrm{Ly}\alpha,\,\mathrm{Max}}$ and the maximum of the read peak matches. Thus, in this method, we compute the Ly α wavelength as

$$\lambda_{\mathrm{Ly}\alpha,\mathrm{obs}} = \lambda_{\mathrm{Ly}\alpha,\mathrm{Max}} - \mathrm{FWHM}_{\mathrm{Red}}, \tag{7}$$

where $\mathrm{FWHM}_{\mathrm{Red}}$ is the FWHM of the red peak of the Ly α line profile. This relation is compatible with the observational results found in Verhamme et al. (2018), who first suggested this kind of correction. However, this trend depends strongly in the outflow geometry that is assumed (Gurung-López et al. 2019b). In particular, this relation works well for the Thin Shell geometry, while the Galactic Wind deviates slightly from it. In this work, we use equation (7) for both the Thin Shell and the Galactic Wind in order to qualitatively study the impact in the clustering of using a relation that is slightly off. This mimics the observational framework in which an outflow geometry is assumed for all the LAE population, while the real escape channel has some differences with the assumed model.

In general, the GM, IC, and GM-F algorithms, as presented here, are biased estimators of the redshift. For example, as the maximum of the Ly α line is usually redder than Ly α, GM and IC usually provide $\lambda_{\mathrm{Ly}\alpha,\,\mathrm{obs}} > \lambda_{\mathrm{Ly}\alpha}$. Therefore, the distribution of $\lambda_{\mathrm{Ly}\alpha,\,\mathrm{obs}}$ is not centred at $\lambda_{\mathrm{Ly}\alpha}$. Also, in the case of the GM-F algorithm, the

relation between the FWHM and the offset of the global maximum might change depending on redshift and outflow geometry, leading to a biased estimation. In practice, these systematic biases can be corrected. For example, in Steidel et al. (2010) and Rudie et al. (2012), authors compared the redshift provided by the IC method and the redshift provided by other spectral features, such as H α. Then, they used the systematic offset between these two distributions as a correction for the redshift inferred only from Ly α.

In the following, we leave the GM, IC, and GM-F methodologies uncorrected by systematic biases. In this way, it becomes more apparent the reason behind some of the trends that we find (see Section 5). This choice does not affect the recovered clustering, as it is insensitive to the mean of the $\lambda_{\mathrm{Ly}\,\alpha,\,\mathrm{obs}}$ distribution (e.g. Byrohl et al. 2019).

4.2 Neural networks

We propose a novel method to determine the systemic redshift of LAEs from their whole Ly α line profile through a neural network. In this section, we explain all the ingredients of the neural networks implemented in this work. First, we describe the architecture and the different training sample. Then, we study how the full line profile helps us determine the Ly α wavelength.

4.2.1 Neural network architecture

The architecture of the neural network in this work consists in an input layer, a single hidden layer, and an output layer. We remark that this work does not focus on finding the best architecture to solve the Ly α wavelength determination problem, as this would depend on the Ly α observation characteristics (e.g. spectral resolution, signal-to-noise ratio, etc.). Instead, we adopt this simple architecture as a proof of concept.

In this work, we seek for an algorithm that could be replicated in observational experiments. With this goal in mind, for a given Ly α line profile $\phi(\lambda)$, we set as input $\phi(\Delta\lambda_{\mathrm{Obs}})$, where we have mapped

$$\lambda \rightarrow \Delta\lambda_{\mathrm{Obs}} = \lambda - \lambda_{\mathrm{Ly}\,\alpha,\mathrm{Max}}. \tag{8}$$

In this way, the global maximum of the Ly α line profile is always centred at $\Delta\lambda_{\mathrm{Obs}} = 0$, which can be easily replicated in observational experiments. Additionally, we rescale each individual $\phi(\Delta\lambda_{\mathrm{Obs}})$ in a way that the minimum of the line profile is 0 and the maximum is 1.

4.2.2 Training sets

Throughout this work, we implement two neural networks. Both of them have the same architecture, but different training sets:

(i) **NN:Uniform**: The training sample is randomly selected from the whole LAE population. In other words, there is no dependence on any LAE property.

(ii) **NN:Bright**: The training sample is constructed only by the brightest LAEs. In practice, we rank our LAE population by their Ly α luminosity. Then, we split the LAE population in two and use for training the brightest subset.

The motivation behind each training set is different. On one hand, NN:Uniform represents an ideal scenario where, in a given survey, a subset of the whole LAE population is homogeneously selected and re-observed at a wavelength range that allows the measurement of spectral features (other than the Ly α line) and the assignment of the true systemic rest frame. However, it is, in general, challenging

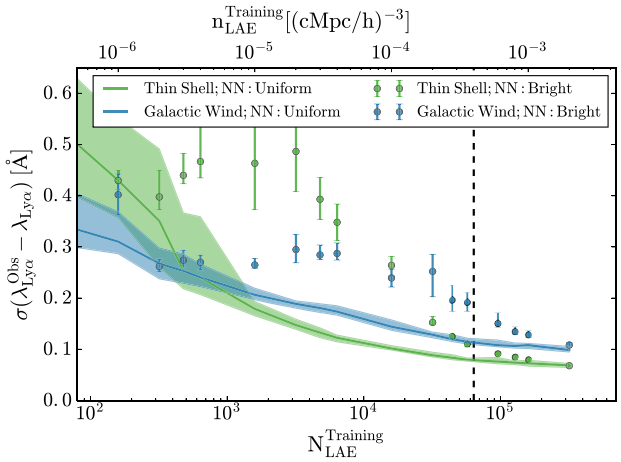

Figure 4. Comparison of the accuracy of the neural network as a function of the training sample size at redshift $z = 3.0$. The solid lines show the NN:Uniform algorithm, while the dots show the NN:Bright. The shaded regions and error bars are the 1σ dispersion for the NN:Uniform and NN:Bright algorithms, respectively. The Thin Shell is represented in green while the Galactic Wind is plotted in blue. The black dashed line shows the fiducial cut in number density for the training sample adopted through this work, which corresponds to a 10 per cent of the number density of the full LAE sample.

to obtain systemic redshifts of such a homogeneous subsample, since the flux of other spectral features tends to be less prominent than Ly α (Trainor et al. 2015). On the other hand, NN:Bright is designed to study how well the LAE redshift determination works even only with the brightest LAEs re-observed at other wavelength, which is closer to a realistic situation than the NN:Uniform case.

We train both neural networks (NN:Uniform and NN:Bright) for each combination of redshift, outflow geometry, and spectral quality (see Section 5).

The performance of the neural networks is linked to the size of the training set. In general, the larger the training set, the more accurate the neural network becomes. In Fig. 4, we show the accuracy of our neural networks (NN:Uniform and NN:Bright) for the different outflow geometries at redshift 3.0 as a function of the number density of LAEs used for the training, $n_{\mathrm{LAE}}^{\mathrm{Training}}$, and the training set size $N_{\mathrm{LAE}}^{\mathrm{Training}}$. Here, we use the standard deviation of $\lambda_{\mathrm{Ly}\,\alpha}^{\mathrm{Obs}} - \lambda_{\mathrm{Ly}\,\alpha}$ (noted as $\sigma(\Delta\lambda)$) to quantify the quality of the neural network, as the clustering of LAEs is sensitive to the distribution of $\Delta\lambda$ (Byrohl et al. 2019). We remark that the mean and median of $\lambda_{\mathrm{Ly}\,\alpha}^{\mathrm{Obs}} - \lambda_{\mathrm{Ly}\,\alpha}$ are, in general, one order of magnitude smaller than $\sigma(\Delta\lambda)$, and they have little impact on the two-point statistics of the autocorrelation function. In order to estimate the variance of $\sigma(\Delta\lambda)$, for each value of $N_{\mathrm{LAE}}^{\mathrm{Training}}$ we performed 100 iterations if $n_{\mathrm{LAE}}^{\mathrm{Training}} < 4 \times 10^{-4}$ $(\mathrm{cMpc}\,h)^{-3}$ and 10 iteration if $n_{\mathrm{LAE}}^{\mathrm{Training}} > 4 \times 10^{-4}$ $(\mathrm{cMpc}\,h)^{-3}$.

Throughout this work, we use rest-frame length units to quantify $\Delta\lambda$. However, it is common in the literature to provide this quantity in velocity units (e.g. Byrohl et al. 2019). These two quantities are equivalent and can be transformed as $\Delta v = c\Delta\lambda/\lambda_{\mathrm{Ly}\,\alpha} \simeq (247\ \mathrm{km\,s^{-1}}) \times \Delta\lambda/1$ Å.

As we show in Fig. 4, the neural networks are able to assign a $\lambda_{\mathrm{Ly}\,\alpha}^{\mathrm{Obs}}$ close to $\lambda_{\mathrm{Ly}\,\alpha}$. Overall, the accuracy increases as the training sample size is increased. However, the NN:Uniform and NN:Bright algorithms behave slightly different. On one hand, in the NN:Uniform

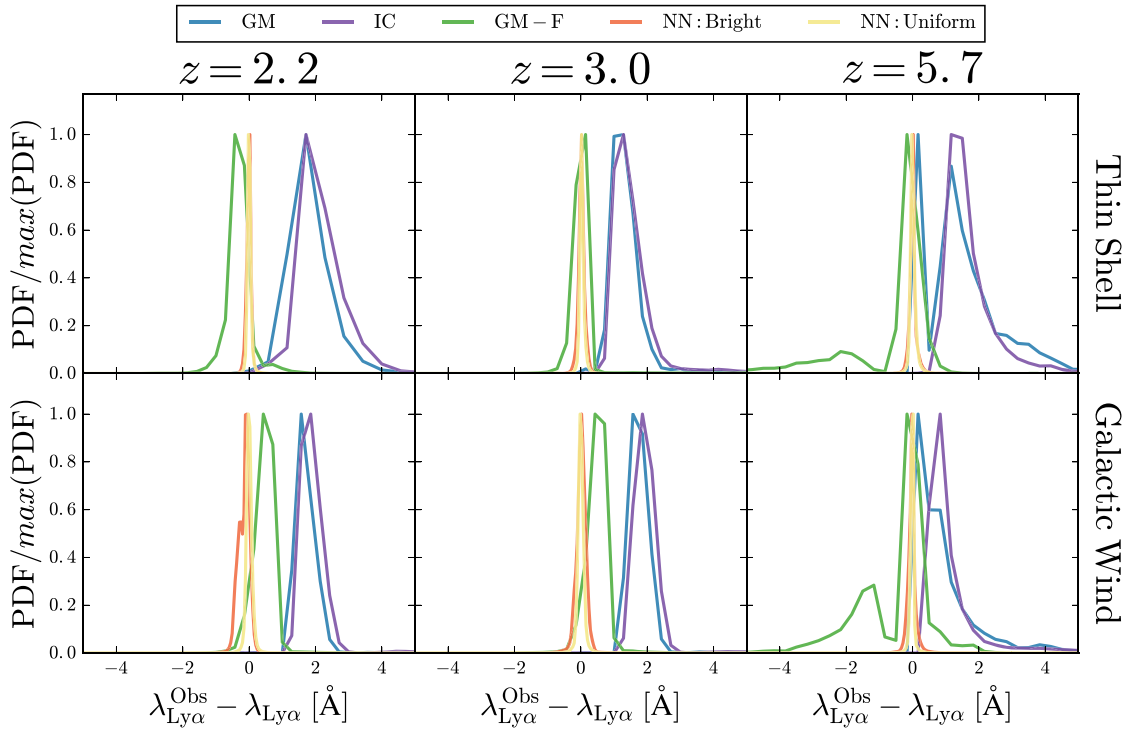

Figure 5. Distribution of the difference between the assigned Ly α wavelength and the intrinsic $\lambda_{{\rm Ly}\,\alpha}$ for each of the different Ly α identification algorithms. The GM algorithm is displayed in blue, IC in purple, GM-F in green, NN:Bright in orange, and NN:Uniform in yellow. Each column shows a different redshift bin (2.2, 3.0, and 5.7 from left to right). The models using the Thin Shell (Galactic Wind) geometry are shown in the top (bottom) panels.

algorithms, $\sigma(\Delta\lambda)$ decreases until $n^{\rm Training}_{\rm LAE} \sim 4 \times 10^{-4}$ $({\rm cMpc}\,h)^{-3}$ ($\sim$ 10 per cent of the total LAE sample), where it reaches a plateau around 0.1 Å. This means that at this value of $n^{\rm Training}_{\rm LAE}$, the training sample is big enough to cover the full variety of Ly α line profiles. Thus, adding more galaxies to the training sample beyond 10^{-4} $({\rm cMpc}\,h)^{-3}$ add little information, leaving the accuracy constant. On the other hand, the performance of the NN:Bright algorithm is worse at low $n^{\rm Training}_{\rm LAE}$. Meanwhile, the NN:Bright converges to the NN:Uniform accuracy when $n^{\rm Training}_{\rm LAE} \sim 2 \times 10^{-3}$ $({\rm cMpc}\,h)^{-3}$ ($\sim$50 per cent of the total sample). This is because the Ly α line profile depends on galaxy and IGM properties (see Figs 2 and 3). In particular, the Ly α line profile of the bright LAEs is more redshifted than the Ly α line profiles of the faint LAEs. Hence, the training sample in the NN:Bright is biased towards redshifted Ly α line profiles and it does not contain typical line profiles from faint LAEs. This reduces the accuracy at low $n^{\rm Training}_{\rm LAE}$ in comparison to NN:Uniform, which makes an uniform selection on $L_{{\rm Ly}\,\alpha}$. Then, for larger values of $n^{\rm Training}_{\rm LAE}$ the training samples of the NN:Uniform and NN:Bright become more similar, which makes them converge to the same accuracy. Finally, the line profiles generated using the Galactic Wind outflow geometry seem slightly more complex than the Thin Shell counterparts, which results, typically, in a better accuracy for the Thin Shell geometry.

From now on, we fix the number density of the training sample to $n^{\rm Training}_{\rm LAE} = 4 \times 10^{-4}$ $({\rm cMpc}\,h)^{-3}$. We have chosen this value of $n^{\rm Training}_{\rm LAE}$ for two main reasons: (i) the information of the training sample of the NN:Uniform saturates and increasing $n^{\rm Training}_{\rm LAE}$ does not add new information to it and (ii) the NN:Bright has not yet converged to the NN:Uniform accuracy, so we can study the differences between these two methodologies.

5 THE EFFECTS OF THE Ly α WAVELENGTH DETERMINATION IN IDEAL LINE PROFILES

Hereafter, we study how the misidentification of the Ly α wavelength modifies the clustering. In order to understand the physical consequences, in this section, we rather focus on an ideal case where we ignore binning artefacts and the instrumental noise in a LAE spectrum. We will consider more realistic situations in next section.

5.1 Algorithm performances in ideal Ly α line profiles

In this section, we compare the performance of the four methodologies to determine $\lambda_{{\rm Ly}\,\alpha}$ from an Ly α line profile. In Fig. 5, we show the probability distribution function (PDF) of the deviation of $\lambda^{\rm Obs}_{{\rm Ly}\,\alpha}$ from $\lambda_{{\rm Ly}\,\alpha}$ ($\Delta\lambda$) for the different redshifts, outflow geometries, and algorithms to determine $\lambda^{\rm Obs}_{{\rm Ly}\,\alpha}$. Also, we list the mean ($\mu(\Delta\lambda)$) and standard deviation ($\sigma(\Delta\lambda)$) of these distribution in Table 1. Overall, the algorithms using neural networks outperform the standard algorithms (GM, IC, and GM-F). We find that the best methodology to retrieve $\lambda_{{\rm Ly}\,\alpha}$ is NN:Uniform (yellow), as the standard deviation of $\Delta\lambda$ is the smallest at all redshifts and outflow geometries. In detail, for all our models using NN:Uniform, $\sigma(\Delta\lambda)$ is below 0.1 Å and $\mu(\Delta\lambda)$ is lower than 0.01 Å. The NN:Uniform is followed closely by the NN:Bright(orange), which also exhibits a great performance, with $\sigma(\Delta\lambda) \sim 0.15$ and $\mu(\Delta\lambda) < 0.01$ Å.

Regarding the standard methodologies, in general, GM-F performs better than GM. Meanwhile, IC is the methodology with the worst performance at $z = 2.2$ and $z = 3.0$, while it performs better than GM and GM-F at $z = 5.7$. If we focus on GM (blue) and IC (purple), we find that GM presents a smaller dispersion than IC at redshifts 2.2 and 3.0, while the opposite is true at $z = 5.7$. We also find that the

Table 1. Mean ($\mu(\Delta\lambda)$) and standard deviation ($\sigma(\Delta\lambda)$) of the difference between the Ly α assigned as Ly α and the true Ly α frequency for the different Ly α identification algorithms, redshifts, and outflow geometries.

Redshift	Geometry	Algorithm	μ (Å)	σ (Å)
2.2	Thin Shell	GM	1.822	0.66
		IC	2.1634	0.72
		GM-F	− 0.3059	0.45
		NN:Uniform	0.0032	0.07
		NN:Bright	− 0.0087	0.08
	Galactic Wind	GM	1.8289	0.89
		IC	2.0189	0.94
		GM-F	0.4394	0.36
		NN:Uniform	0.0022	0.11
		NN:Bright	− 0.156	0.17
3.0	Thin Shell	GM	1.4747	0.99
		IC	1.6347	1.07
		GM-F	0.0251	0.52
		NN:Uniform	0.0303	0.08
		NN:Bright	0.0202	0.1
	Galactic Wind	GM	1.9134	0.99
		IC	2.1233	1.13
		GM-F	0.4757	0.34
		NN:Uniform	0.0007	0.11
		NN:Bright	0.0112	0.16
5.7	Thin Shell	GM	1.5134	1.22
		IC	1.7079	0.8
		GM-F	− 0.7141	1.44
		NN:Uniform	0.0116	0.1
		NN:Bright	0.0094	0.12
	Galactic Wind	GM	0.8751	0.97
		IC	1.131	0.85
		GM-F	− 0.5247	1.09
		NN:Uniform	− 0.0132	0.09
		NN:Bright	− 0.0185	0.17

performances of GM and IC are similar for both outflow geometries; the mean of the $\Delta\lambda$ distribution is shifted ∼1.7 Å redwards $\lambda_{\mathrm{Ly}\,\alpha}$. This is a direct consequence of the Ly α RT, as the photons get redshifted when they travel through the neutral hydrogen of outflows. Then, as GM takes the global maximum of the Ly α line profile as $\lambda^{\mathrm{Obs}}_{\mathrm{Ly}\,\alpha}$, the whole $\Delta\lambda$ is systematically redshifted. In fact, we find that for the GM algorithm, the peak of the $\Delta\lambda$ distribution is located at the same position than the peak of their corresponding stacked line profiles (see Fig. 1). Moreover, for the Thin Shell at $z = 5.7$ the $\Delta\lambda$ distribution of the GM algorithm exhibits a double peak shape, and the same position than the peaks present in the Ly α stacked line profile of that model.

The performance of GM-F (green) depends strongly on the outflow geometry and redshift. On one hand, at low redshift ($z = 2.2$ and $z = 3.0$), GM-F performs better than GM and IC in both outflow geometries, as the GM-F distributions of $\Delta\lambda$ are thinner than those of GM and IC. However, the GM-F performance in the Thin Shell is better than in the Galactic Wind. In particular, $\mu(\Delta\lambda)$ is close to zero for the Thin Shell, while it is ∼0.5 Å in the Galactic Wind. This is a consequence of the different relation between $\lambda_{\mathrm{Ly}\,\alpha,\,\mathrm{Max}}$ and $\mathrm{FWHM_{Red}}$ in the Thin Shell and Galactic Wind outflow geometries (see fig. 4 in Gurung-López et al. 2019b). On the other hand, at $z = 5.7$ the $\Delta\lambda$ distribution becomes bimodal, with a second less prominent peak centred around $\Delta\lambda = -2$ Å (see Fig. 1). This is caused by the IGM modifying the Ly α line profile in such a way that the $\mathrm{FWHM_{Red}}$ of the observed Ly α line profile is larger than the initial.[1] As a consequence, GM-F overcorrects the shift of the Ly α peak, causing the peak in the $\Delta\lambda$ distribution at $\Delta\lambda < 0$.

5.2 Impact on the redshift-space clustering of LAEs

The misidentification of the Ly α wavelength has a non-negligible impact on the three-dimensional clustering of LAE samples that rely only on their Ly α line profile to determine their redshift, and hence, their radial position (Byrohl et al. 2019). Indeed, the measured redshift of LAEs has three main contributions: (i) the geometric redshift given by the Hubble flow, (ii) the redshift or blueshift given by the peculiar velocity of the galaxy along the line of sight (Kaiser 1987), normally dabbed RSD, and (iii) a redshift or blueshift rising from the Ly α wavelength misidentification, i.e. $\Delta\lambda \neq 0$.

In order to characterize the clustering of a galaxy population, it is useful to define the galaxy overdensity field

$$\delta_{\mathrm{g}} = \frac{n_{\mathrm{g}}(\boldsymbol{x})}{\langle n_{\mathrm{g}}\rangle} - 1, \tag{9}$$

where $n_g(\boldsymbol{x})$ is the number density galaxies at the position $\boldsymbol{x}$, and $\langle n_{\mathrm{g}}\rangle$ is its average value. Then, the two-point correlation function (2PCF), ξ_{g}, is defined as

$$1 + \xi(\boldsymbol{r}) = \langle [1 + \delta_{\mathrm{g}}(\boldsymbol{x})][1 + \delta_{\mathrm{g}}(\boldsymbol{x} + \boldsymbol{r})]\rangle, \tag{10}$$

where $\boldsymbol{r}$ is the pair vector between two points separated a distance r. We also consider the power spectrum $P_{\mathrm{g}}(\boldsymbol{k})$, which is the Fourier transform of the 2PCF, i.e.

$$P_{\mathrm{g}}(\boldsymbol{k})\,(2\pi)^3\,\delta_D(\boldsymbol{k} + \boldsymbol{k'}) = \langle \delta_{\mathrm{g}}(\boldsymbol{k})\,\delta_{\mathrm{g}}(\boldsymbol{k'})\rangle, \tag{11}$$

where $\boldsymbol{k}$ is the wavenumber, and $\delta_D(\boldsymbol{k})$ is the Dirac delta function.

Throughout this work, we will focus on the clustering in *redshift space*. To incorporate the three redshift contributions to our clustering analysis, we recompute the position of our LAEs in redshift space as

$$\boldsymbol{s} = \boldsymbol{r} + \frac{V_{\mathrm{LoS}} + \Delta V_{\mathrm{Ly}\alpha}}{a(z)H(z)}\hat{Z}, \tag{12}$$

where we take Z as the direction of the line of sight, assuming the global plane-parallel approximation (Beutler et al. 2014), V_{LoS} is the velocity along the line of sight of the galaxy. $a(z)$ and $H(z)$ are the scale factor and the Hubble parameter at redshift z, respectively. Additionally,

$$\Delta V_{\mathrm{Ly}\,\alpha} = c\left(1 - \frac{\lambda_{\mathrm{Ly}\,\alpha}}{\lambda^{\mathrm{Obs}}_{\mathrm{Ly}\,\alpha}}\right), \tag{13}$$

where c is the speed of light. Finally, there is a fraction of the LAE population that are shifted outside of the box after transforming their line position to redshift space, i.e. including the contributions of V_{LoS} and $\Delta V_{\mathrm{Ly}\,\alpha}$. For these galaxies, we assume that our simulation box is periodic along the line of sight. Therefore, galaxies with $\pi < 0$, they are assigned $\pi = L_{\mathrm{Box}} + \pi$ and galaxies with $\pi > L_{\mathrm{Box}}$, they are assigned $\pi = \pi - L_{\mathrm{Box}}$.

5.2.1 *Clustering damping in Fourier space*

First, we focus on the impact of the Ly α misidentification in the power spectrum. In Byrohl et al. (2019), the authors analytically showed that, due to the Ly α misidentification, the amplitude of

[1] For an example of this, see fig. 7 of Gurung-López et al. (2020).

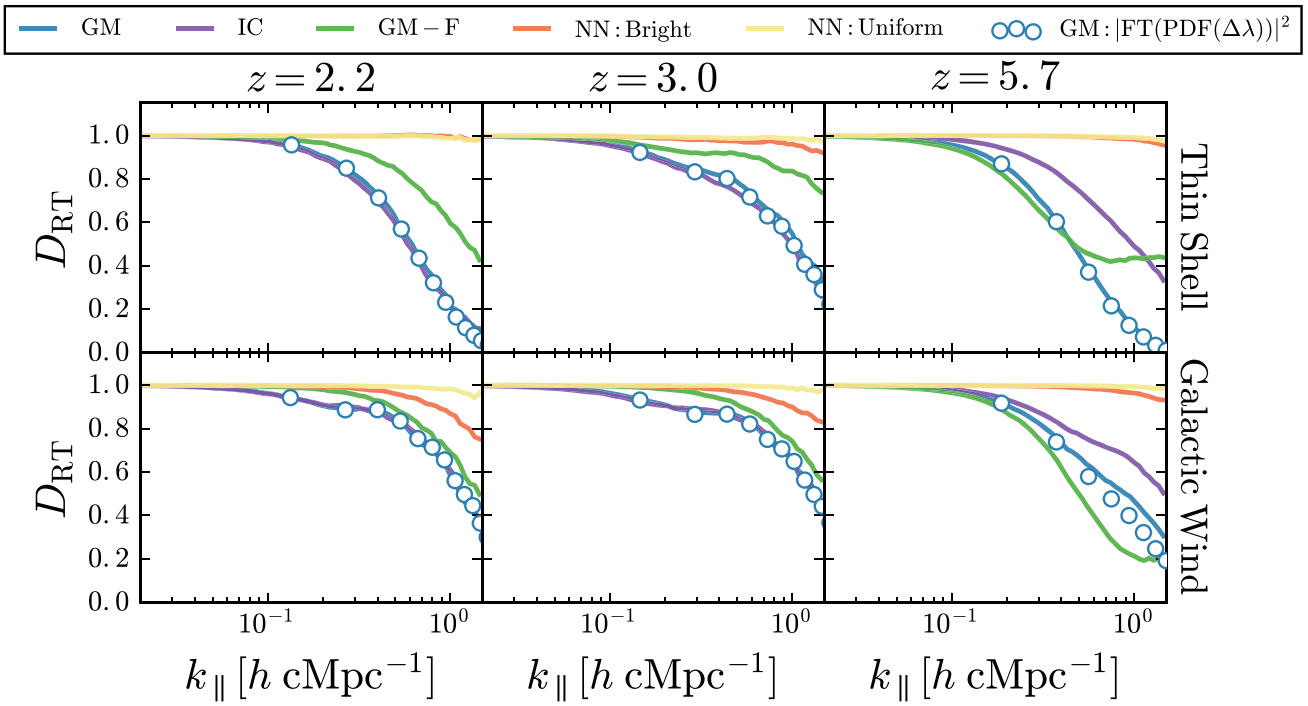

Figure 6. Damping of the power spectrum along the line of sight at redshifts 2.2, 3.0, and 5.7 from left to right. In the top panels, we show the models with the Thin Shell and Galactic Wind outflow geometry in the top and lower panels, respectively. The coloured solid lines show the different algorithms to identify the Ly α wavelength (the colour code is the same as in Fig. 5). The blue empty dots display the damping of the power spectrum computed from the PDF of $\Delta\lambda$ through equation (14).

the power spectrum is reduced along the line of sight at large wavenumber ($k_{\shortparallel}$) values, i.e. at small scales. We define the damping of the power spectrum due to the Ly α misidentification as in Byrohl et al. (2019), i.e.

$$D_{\rm RT}(k_{\parallel}) = \frac{P^{\rm RT}(k)}{P(k)} = |FT(PDF(\Delta\lambda))|^2, \quad (14)$$

where $P(k_{\shortparallel})$ is the intrinsic redshift-space power spectrum of the LAE population, i.e. setting $\Delta V_{{\rm Ly}\,\alpha} = 0$, in equation (12). Also, $P^{\rm RT}(k_{\shortparallel})$ is the power spectrum of the LAE sample after including the displacement along the line of sight due to the Ly α misidentification. Finally, the Fourier transformation is indicated as *FT*. We estimate the power spectrum from the simulated LAEs by making use of the Fast Fourier Transformation (FFT). We set the number of grids as 512^3 with which the Nyquist wavenumber is $k_{\rm Nyq} \sim 3\,h{\rm cMpc}^{-1}$. The last equality holds only if (i) the moments of PDF($\Delta\lambda$) are scale independent and (ii) $\Delta\lambda$ is uncorrelated with the large-scale density and velocity fields (Byrohl et al. 2019).

In Fig. 6, we show the damping in the power spectrum for our different Ly α wavelength recovering algorithms and multiple models. In general, we find that the amplitude of the power spectrum including the Ly α misidentification is lower than the intrinsic power spectrum at the scales relevant to the BAO and RSD measurements, as $D_{\rm RT} < 1$ at $0.1 < k_{\parallel}\,[h\,{\rm cMpc}^{-1}] < 1$. In particular, the impact is greater on smaller scales (larger $k_{\shortparallel}$), while at large enough scales it disappears. This damping can be interpreted as the Finger-of-God effect and matches the results from Byrohl et al. (2019), although the detailed suppressions behave differently as the PDFs are different.

$D_{\rm RT}$ has been computed in two different ways: (i) computing the power spectra directly from the LAE positions in our simulation box (solid lines) and (ii) by computing the Fourier transform of the one point PDF of $\Delta\lambda$ (open circle points). We compare the two methods only for the GM in Fig. 6, and confirm that they are in a good agreement. This ensures that the misidentification is uncorrelated with the large-scale density or velocity field, while there is a small hint of the IGM interaction at $k_{\shortparallel} \sim 1$ for the Galactic Wind at $z = 5.7$. We have checked similar results for the other algorithms, and hence omitted them in the figure.

We find that the algorithms with a higher accuracy for recovering the Ly α wavelength from the line profile show a shallower damping of the power spectrum. In particular, the NN:Uniform is the algorithm that is the least affected by the Ly α wavelength misidentification. In fact, the recovered power spectrum agrees at the 1 per cent level up to $k_{\shortparallel} = 1$ for both outflow geometries and at all redshifts. The second best performance is achieved by the NN:Bright, which exhibits up to ∼0.2 decreases in the power spectrum amplitude in the Galactic Wind, while in the Thin Shell, the damping is slight stronger than in the NN:Uniform. Then, the GM, IC, and GM-F algorithms are heavily affected by the Ly α misidentification, as the amplitude of the power spectrum decreases dramatically on small scales at all redshifts and outflow geometries. In particular, GM-F is less affected than GM and IC, at redshifts 2.2 and 3.0. Meanwhile at $z = 5.7$, IC behaves better than the other two and the performance of GM-F is comparable (in the Thin Shell) or slight worst (in the Galactic Wind) than GM.

5.2.2 *Impact on the 2D clustering in configuration space*

In this section, we explore the 2PCF to illustrate qualitatively the clustering distortion produced by the misidentification of the Ly α wavelength. Later, we will further quantify the anisotropic distortions using the Legendre multipole moments.

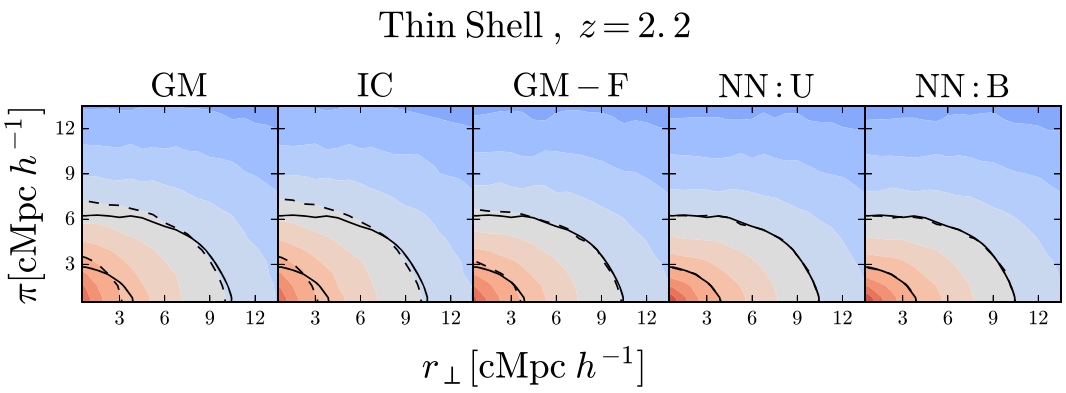

Figure 7. Redshift-space clustering divided the parallel (π) and perpendicular ($r_\perp$) directions to the line of sight at $z = 2.2$ for the Thin Shell outflow geometry. From left to right, in each panel, we display the GM, IC, GM-F, NN:Uniform, and NN:Bright algorithms, respectively. For each algorithm, we show the clustering levels $\xi(r_\perp, \pi) = 10^{-0.6}$ and $10^{0.0}$ in dashed black lines in their corresponding panels. The solid black lines are the same in every panel and correspond with the same clustering levels in the case in which there is no Ly α wavelength misidentification.

We estimate 2PCF with the standard Landy–Szalay estimator (Landy & Szalay 1993). Fig. 7 shows the clustering divided into parallel and perpendicular to the line-of-sight components for our different algorithms. In particular, we display the model at redshift 2.2 using the Thin Shell geometry, but similar results are found for the other models too. To compare the performance of the algorithms, we show the contours with the clustering amplitude of $\xi(r_\perp, \pi) = 10^{-0.6}$ and $10^{0.0}$ for no Ly α wavelength misidentification (solid) and for each algorithm (dashed). Overall, the misidentification of the Ly α wavelength causes an elongation of the LAE clustering along the line of sight. This elongation is more prominent for the algorithms with worst performance recovering the Ly α wavelength. In concordance with our previous findings, the GM, IC, and GM-F algorithm fail to recover the intrinsic redshift-space clustering of LAEs. Meanwhile, NN:Uniform and NN:Bright achieve almost a perfect recovery of the 2PCF.

5.2.3 *Impact on the monopole*

The multipole 2PCF is given by

$$\xi_\ell(s) = \frac{2\ell + 1}{2} \int_{-1}^{1} \mathrm{d}\mu \; \xi(s, \mu) \, \mathscr{L}_\ell(\mu), \tag{15}$$

where ℓ is the multipole degree, $\mathscr{L}_\ell$ is the Legendre's polynomial of degree ℓ, and μ is the cosine between the line of sight and the separation vector of a galaxy pair.

In Fig. 8, we show the ratio of ξ_0, the monopole ($\ell = 0$) of our LAE populations by using our different algorithms to determine the Ly α wavelength, and ξ_R, the monopole when the Ly α wavelength is identified perfectly, i.e. $\Delta\lambda = 0$ for every galaxy. Overall, we find that the method used to determine the Ly α wavelength change the retrieved monopole of LAEs at scales below 10 cMpc h^{-1} for both outflow geometries at all redshifts. Meanwhile, the large-scale clustering (> 10 cMpc h^{-1}) remains unchanged. In particular, at large scales, the monopole of the different algorithms converged to the intrinsic one with nearly no difference among them.

The clustering for the traditional algorithms (GM, IC, and GM-F) is suppressed at small scales. In contrast, the measured monopole using both our neural networks (NN:Uniformand NN:Bright) match extraordinary well the intrinsic clustering of LAEs at all scales. These differences in the monopole between the standard approaches and the neural networks are driven by the much better performance of NN:Uniform and NN:Bright when determining the Ly α frequency (see Fig. 5). In particular, the large dispersion of $\Delta\lambda$ given by GM, IC, and GM-F translates in to a large scatter of $\Delta V_{\mathrm{Ly}\,\alpha}$. Which means that, the position of the galaxies in redshift space is quite spread along the line of sight with respect to their original position in redshift space. This dilutes the clustering on small scales along the line of sight, which causes a decrease of power in the monopole on scales $\gtrsim 1$ cMpc h^{-1}.

Moreover, we find that, on the scales studied here, both neural networks produce very similar monopoles at all redshifts and for both outflow geometries. Meanwhile, we see that the power suppression in the GM-F monopole is generally smaller than the suppression in the GM monopole. Again, this comes from the different performance among these algorithms. GM-F performs better than GM since it provides a tighter distribution of $\Delta\lambda$ in the PDF (Fig. 6).

5.2.4 *Impact on the quadrupole*

The quadrupole ($\ell = 2$) is more sensitive to the anisotropic clustering than the monopole. In Fig. 9, we show the quadrupole at different redshifts for both outflow geometries. The solid lines indicate the quadrupole when the peculiar motion of galaxies and the Ly α wavelength misidentification are implemented (using equation 12). In contrast, in order to isolate the contribution of the Ly α shift, we show with dashed lines the quadrupole using equation (12) but assuming that galaxies have no peculiar motion along the line of sight ($V_\mathrm{Los} = 0$).

Overall, we find that the uncertainty in the Ly α wavelength determination changes the ratio between the clustering parallel and perpendicular to the line of sight. In this way, the quadrupole amplitude is enhanced at scales $\lesssim 10$ cMpc h^{-1}. This result is consistent with the Finger-of-God effect as we have already confirmed in Fourier space. The Ly α misidentification can be regarded as an additional random shuffling along the line of sight. This interpretation is further assured by the fact that the dashed lines have negligible quadrupole amplitudes on large scales, $\gtrsim 10$ cMpc h^{-1}. Additionally, all methodologies converge on large scales, $\gtrsim 10$ cMpc h^{-1}, suggesting that the quadrupole 2PCF on such scales can be safely used to infer the peculiar velocity contribution in LAE surveys. Finally, the negative amplitude of the quadrupole 2PCF at large scales is qualitatively consistent with the Kaiser effect.

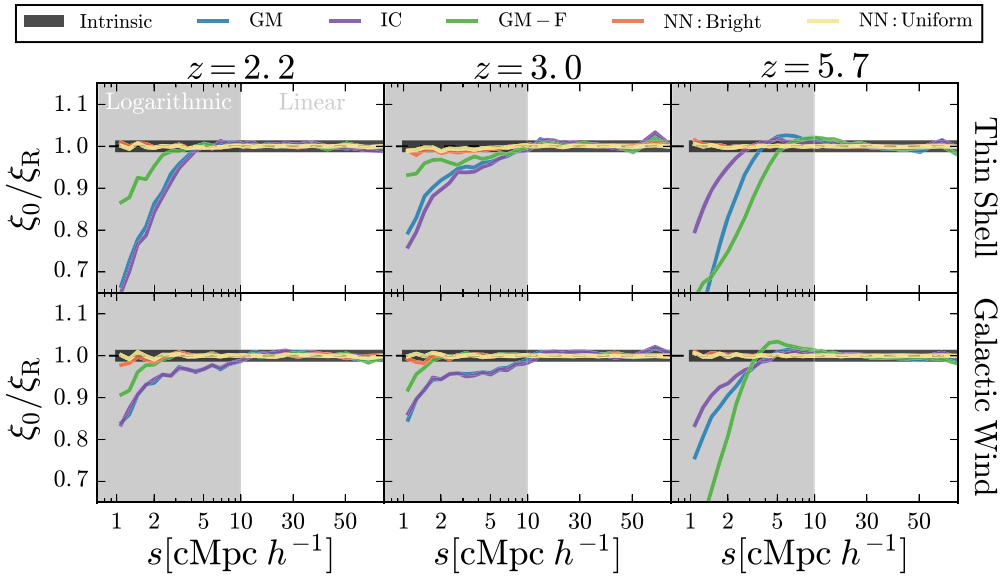

Figure 8. Ratio between the monopole in redshift space of the LAEs samples using different Ly α frequency identification algorithms (thin coloured lines) and the one assuming a perfect accuracy in the Ly α wavelength identification (thick black line). Each column displays a redshift bin (2.2, 3.0, and 5.7 from left to right). Top (bottom) panels display the Thin Shell (Galactic Wind) outflow geometry models. The scale in the grey shaded region is logarithmic, while in the white region it is linear.

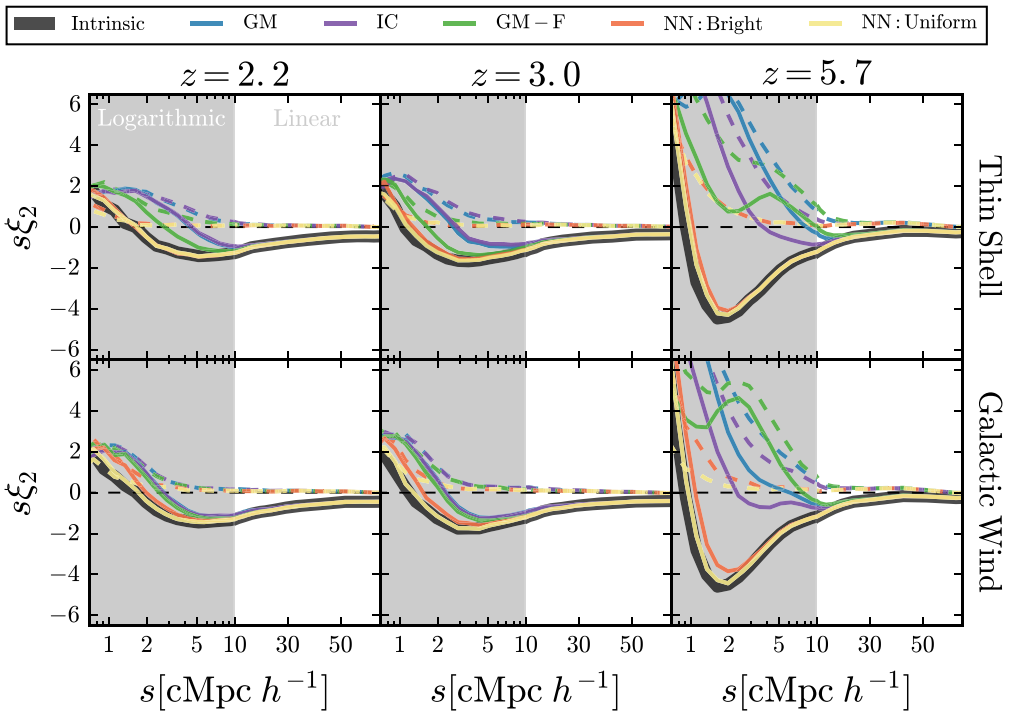

Figure 9. Quadrupole of the LAE samples using Ly α identification algorithms (GM in blue, IC in purple, GM-F in green, NN:Bright in orange, and NN:Uniform in yellow), at different redshifts (2.2, 3.0, and 5.7 from left to right) and implementing different outflow geometries (Thin Shell in the top and Galactic Wind in the bottom). The coloured solid lines are computed with both the contribution of the Ly α misidentification and the peculiar motion of the galaxies. Meanwhile, the coloured dashed line only includes the shift due to the Ly α line profile. The scale in the grey shaded region is logarithmic, while in the white region it is linear.

Focusing on the Ly α wavelength misidentification contribution, we find the same trend as the results in Fourier space. The different algorithms produce different quadrupole predictions, reflecting its algorithms' performance.

On one hand, the quadrupole recovered by the standard algorithms (GM, IC, and GM-F) is heavily distorted with respect the intrinsic one (black). We find that the amplitude of the suppression evolves with redshift, being lighter at low redshift. In fact, the typical scale at

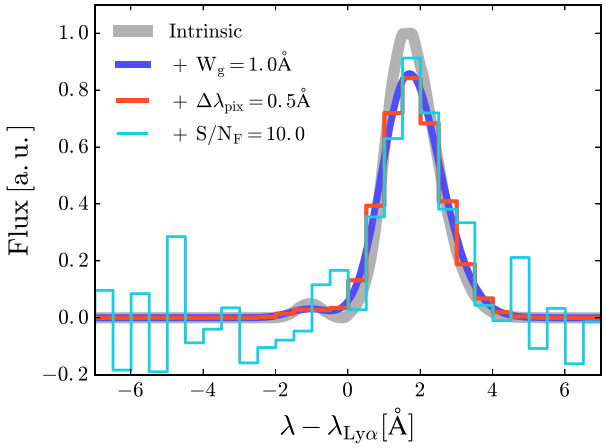

Figure 10. Illustration of the line profile quality downgrade. In grey, we show a particular line profile predicted by our model. The other line displays progressively (and cumulative) lower quality. In blue, we include a Gaussian kernel of FWHM = 1 Å, then we pixelize this line in wavelength bins of size 0.5 Å. Finally, we add Gaussian noise to each pixel with an amplitude so that S/N = 7.

which the intrinsic and the observed LAE quadrupole converge is $\sim$ 5, $\sim$ 8, and $\sim$ 15 cMpc h^{-1} at redshifts 2.2, 3.0, and 5.7, respectively. Additionally, at small scales the quadrupole amplitude is enhanced, and its sign is flipped, specially at redshift 5.7.

Our neural network approaches work pretty well recovering the intrinsic quadrupole at all redshift and for both outflow geometries. In detail, we find very small differences between NN:Uniform and NN:Bright, as NN:Uniform performs slightly better. In other words, the contribution to the quadrupole given by the Ly α wavelength misidentification becomes negligible when the Ly α wavelength is computed using NN:Uniform and NN:Bright.

6 THE EFFECTS OF THE Ly α WAVELENGTH DETERMINATION IN REALISTIC LINE PROFILES.

In the previous sections, we have studied the properties of the Ly α line profiles directly predicted by our model. These Ly α line profiles are ideal in terms of (i) signal to noise, which is effectively infinite and constant across all the Ly α luminosity range of our models (down to $10^{41.5}$ erg s^{-1}), and (ii) the size of the independent wavelength bins (0.1 Å).[2] However, in observational data sets, reaching these conditions is challenging and/or impossible nowadays.

In this section, we study how the quality of the Ly α line profiles affects the clustering measurements in Ly α focused spectroscopic galaxy surveys. We focus on the snapshot at redshift 3.0 using the Thin Shell outflow geometry. We have checked that the same analysis of the other redshift bins and outflow geometries gives, qualitatively, the same trends.

6.1 Mocking measured Ly α line profiles

In the following, we explain how the Ly α line profiles produced by our model are deteriorated with three distinct steps, and we illustrate them in Fig. 10.

[2]This value comes from the bin size used in FLAREON to store the Lyα line profiles.

Step 1) Spectral resolution: We degrade the wavelength resolution of the line profile. For this end, we convolve the Ly α line profile produced by our model with a Gaussian kernel of FWHM W_g. In this way, the line profile gets broader and the features are diluted. For example, the blue peak that is clearly present in the original line profile is hardly seen after a Gaussian filter of 1.0 Å (dark blue solid line in Fig. 10). We implement three different values of W_g: 0.5, 1.0, and 2.0 Å in the LAE's rest frame.

Step 2) Pixelization: We pixelize the Ly α line profile into wavelength bins of width $\Delta\lambda_{pix}$. In practice, the pixelized Ly α line profile ϕ_{pix} is computed as

$$\phi_{pix}(\lambda_{pix}) = \frac{\int_{\lambda_{pix}-\Delta\lambda_{pix}/2}^{\lambda_{pix}+\Delta\lambda_{pix}/2} \phi(\lambda)\, d\lambda}{\Delta\lambda_{pix}}, \quad (16)$$

where λ_{pix} is the wavelength of each wavelength bin. As the width of the wavelength bins increases, it becomes progressively more difficult to resolve features in the Ly α line profile (see the red solid line in Fig. 10). In this work, we implement three different values of $\Delta\lambda_{pix}$: 0.25, 0.5, and 1.0 Å in the LAE's rest frame. For each of these, the full LAE population is convolved with the same value.

Step 3) Noise: Finally, we include noise in the Ly α line profile. In detail, we assign the signal-to-noise ratio of the faintest (in Ly α) object of our catalogue to S/N_F. Then, for each galaxy, its signal to noise ratio is scaled as

$$S/N = S/N_F \times \frac{L_{Ly\alpha}}{L_{Ly\alpha,F}}, \quad (17)$$

where $L_{Ly\alpha}$ is the Ly α luminosity of each galaxy and $L_{Ly\alpha,F}$ is the Ly α luminosity of the faintest galaxy. Note that our models, by construction, provide good estimates of the LAE luminosity function (Gurung-López et al. 2020). Therefore, the S/N distribution should also mimic observations. Next, to each pixel we add Gaussian noise with an amplitude that corresponds to the S/N of that LAE. In this way, LAEs with lower S/N have a noisier Ly α line profile and vice versa. Moreover, the larger S/N_F, the better signal-to-noise ratio has the faintest LAE and the whole LAE population. On the other side, for low S/N_F values, some information from the Ly α line profile (e.g. light blue line) is vanished. In practice, the larger the W_g and $\Delta\lambda_{pix}$, the more information is destroyed for a fixed value of S/N_F. We use four values for S/N_F : 6.0, 7.0, 10.0, and 15.0. Also, for the LAE sample studied here, $L_{Ly\alpha,F} \sim 1.63 \times 10^{42}$ erg s^{-1}.

For each combination of $\{W_g, \Delta\lambda_{pix}, S/N_F\}$, we produce a catalogue of Ly α line profiles. In practice, we could determine directly the Ly α central wavelength from these line profiles, as we did in the ideal case in Section 5. However, as the line profiles are progressively downgraded it becomes more difficult to measure properly the $FWHM_{red}$ (necessary for the GM-F). Also, the global maximum gets more and more discretized as $\Delta\lambda_{pix}$ increases (needed for both, GM, GM-F). In order to alleviate these problems, we fit a Gaussian curve to the most prominent peak of the Ly α line profile and then, measure $FWHM_{red}$ and $\lambda_{Ly\alpha,Max}$ from the Gaussian. Note that, as the line profiles are convolved with a Gaussian kernel of width W_g, $FWHM_{red}$ will be, in general, overestimated, which would translate into a systematic bias for the GM-F method. In order to correct this, we compute the FWHM as

$$FWHM_{red,c} = \sqrt{FWHM_{red}^2 - W_g^2}. \quad (18)$$

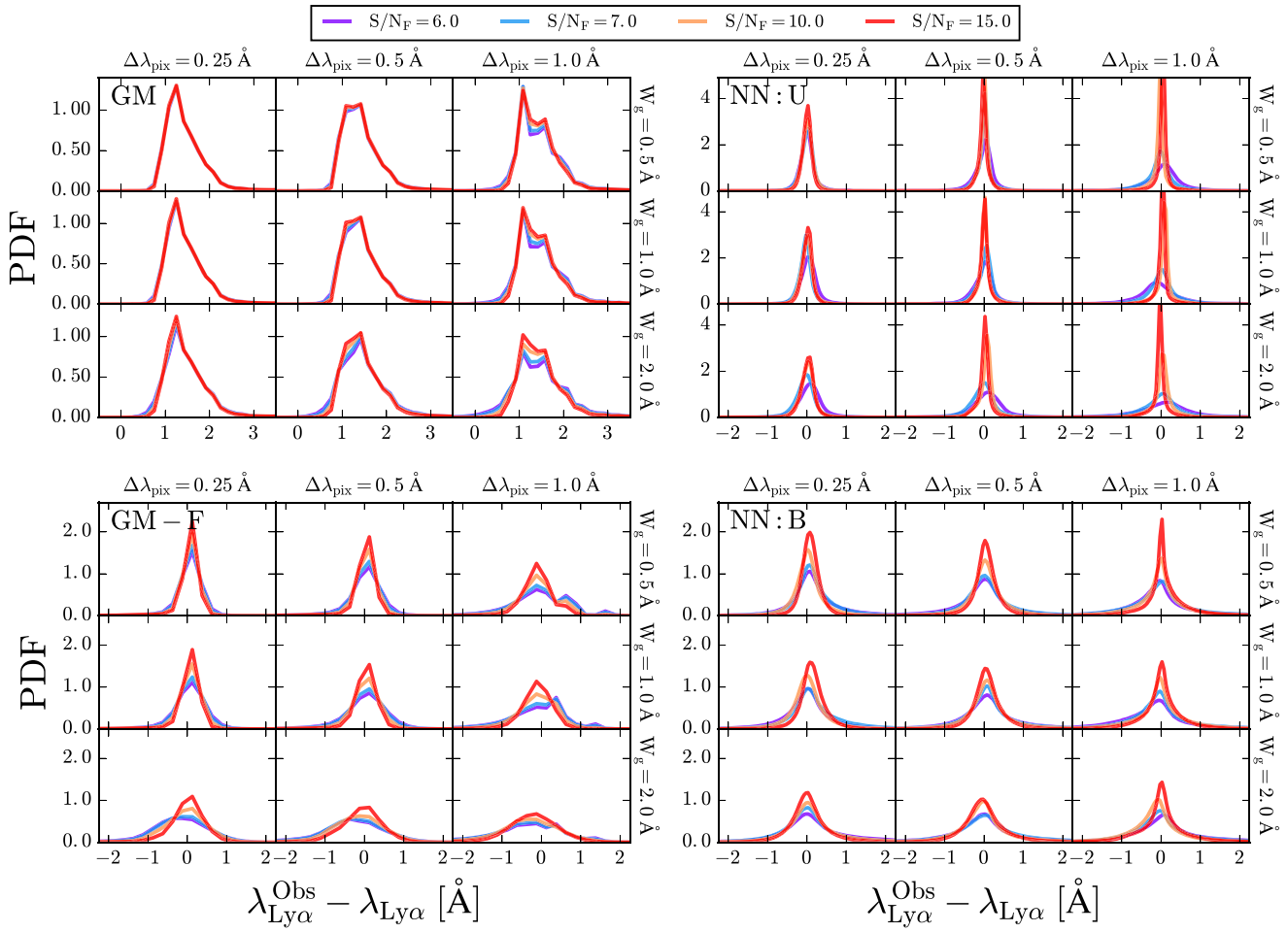

Figure 11. Distribution of the wavelength shift due to the misidentification of the Lyα wavelength from Lyα line profiles for our different spectral qualities and the GM algorithm. The columns show $\Delta\lambda_{pix} = 0.25$, 0.5, and 1.0 Å from left to right. Meanwhile, rows show $W_g = 0.5$, 1.0, and 2.0 Å from top to bottom. Additionally, in each panel, $S/N_F = 6.0$ is displayed in purple, 7.0 in blue, 10.0 in orange, and 15.0 in red. This corresponds to the analysis of the snapshot at redshift 3.0 using the Thin Shell. Top left: GM, bottom left: GM-F, top right: NN:Uniform, bottom right NN:Bright.

There are a few cases [3] where, due to noise in the line profile, $FWHM_{red} < W_g$. In these cases, we assume that $FWHM_{red,c} = FWHM_{red}$. Then, for the standard GM-F algorithm we use $FWHM_{red,c}$ to correct the global maximum offset through equation (7). For completeness, we also show the results for this algorithm without this correction, i.e. using $FWHM_{red}$ instead of $FWHM_{red,c}$. We refer to this last method as uGM-F.

This proceeding can be applied to observational data too. Note that, we only use the Gaussian fitting to compute $FWHM_{red,c}$ and $\lambda_{Ly\alpha,\,Max}$. In this way, the neural networks use directly the modelled line profiles and not the Gaussian resulting from the fitting.

Another consideration is that the performance of the IC algorithm in realistic line profiles depends on more variables than just W_g, $\Delta\lambda_{pix}$, and S/N_F. For example, its accuracy depends on the spectral range considered for computing the centroid of the line profile. If the spectral range is too narrow, then the centroid might be heavily affected by noise. Meanwhile, for a large spectral range the noise might be average out and the IC accuracy would increase. Also, as it was illustrated in Section 5 the behaviour of IC is relatively similar to that of GM, specially at redshifts 2.2 and 3.0. Due to these reasons, we drop the IC algorithm in this section.

[3]The fraction of cases depends on the quality of the line profiles. For example, if $W_g \leq 1.0$ Å and $\Delta\lambda_{pix} \leq 0.5$ Å less than 2 per cent of the sample experience this issue for any S/N_F value. However, when $W_g = 2$ Å and $S/N_F = 6.0$ the percentage increases up to a 25 per cent for $\Delta\lambda_{pix} = 1.0$Å.

6.2 Ly α wavelength displacement

In this section, we compare how the performance of the different algorithms to determine the Lyα wavelength from an Lyα line profile vary with the quality of the spectrum. In Fig. 11, we show the distributions of the displacement between the assigned Lyα wavelength and the true one for the GM, GM-F, NN:Uniform, and NN:Bright algorithms at redshift 3.0 and using the Thin Shell. Then, we sum up the performance of all the redshifts and outflow geometries combinations in Tables 2–7. Overall, for all our four algorithms decreasing the quality of the Lyα line profiles, i.e. decreasing S/N_F and increasing $\Delta\lambda_{pix}$ and W_g, cause a larger uncertainty in the identification of the Lyα wavelength.

Focusing on the GM (Fig. 11 top left), we find that this algorithm is insensitive to lowering the quality of the Lyα line profile. In fact, the mean of the distribution is always centred around 1.5 Å. This is due to the fact that Lyα photos are redshifted as they escape through the

Table 2. Standard deviation (σ) of the difference between the wavelength assigned as Ly α and the true Ly α frequency for different quality configurations and Ly α identification algorithms. This corresponds to the analysis of the snapshot at redshift 3.0 using the Thin Shell.

$z = 3.0$, Thin Shell			GM	GM-F	uGM-F	NN:U	NN:B
W_g^{Rest} (Å)	$\Delta\lambda_{pix}^{Rest}$ (Å)	S/N_F	σ (Å)	σ (Å)	σ (Å)	σ (Å)	σ (Å)
0.5	0.25	6.0	1.33	0.7	0.69	0.52	0.94
		7.0	1.28	0.67	0.66	0.42	0.74
		10.0	1.23	0.61	0.6	0.29	0.49
		15.0	1.21	0.58	0.57	0.2	0.3
	0.5	6.0	1.37	0.83	0.81	0.7	1.24
		7.0	1.34	0.76	0.75	0.53	1.04
		10.0	1.27	0.66	0.65	0.36	0.6
		15.0	1.24	0.61	0.6	0.23	0.41
	1.0	6.0	1.55	1.24	1.2	1.07	1.91
		7.0	1.46	1.07	1.04	0.82	1.5
		10.0	1.32	0.83	0.8	0.46	0.87
		15.0	1.26	0.69	0.67	0.29	0.66
1.0	0.25	6.0	1.31	0.77	0.72	0.57	0.99
		7.0	1.26	0.72	0.68	0.44	0.88
		10.0	1.23	0.66	0.62	0.31	0.56
		15.0	1.18	0.61	0.58	0.21	0.39
	0.5	6.0	1.38	0.93	0.88	0.81	1.43
		7.0	1.33	0.84	0.8	0.61	1.05
		10.0	1.27	0.71	0.67	0.37	0.74
		15.0	1.2	0.64	0.61	0.25	0.46
	1.0	6.0	1.56	1.34	1.32	1.18	2.44
		7.0	1.47	1.14	1.12	0.92	1.57
		10.0	1.34	0.86	0.84	0.51	0.9
		15.0	1.27	0.71	0.69	0.3	0.59
2.0	0.25	6.0	1.29	0.97	0.85	0.79	1.85
		7.0	1.25	0.89	0.77	0.62	1.22
		10.0	1.19	0.78	0.64	0.37	0.88
		15.0	1.12	0.69	0.56	0.26	0.54
	0.5	6.0	1.41	1.22	1.16	1.05	2.21
		7.0	1.36	1.07	0.99	0.81	1.86
		10.0	1.25	0.88	0.76	0.43	0.89
		15.0	1.17	0.76	0.63	0.27	0.62
	1.0	6.0	1.67	1.79	1.76	1.46	2.98
		7.0	1.52	1.53	1.48	1.13	2.22
		10.0	1.35	1.1	1.01	0.61	1.03
		15.0	1.22	0.86	0.73	0.35	0.78

Table 3. Same as Table 2 but displaying the details about the Thin Shell geometry at redshift 2.2.

$z = 2.2$, Thin Shell			GM	GM-F	uGM-F	NN:U	NN:B
W_g^{Rest} (Å)	$\Delta\lambda_{pix}^{Rest}$ (Å)	S/N_F	σ (Å)	σ (Å)	σ (Å)	σ (Å)	σ (Å)
0.5	0.25	6.0	1.57	0.79	0.78	0.67	1.63
		7.0	1.46	0.68	0.67	0.52	1.03
		10.0	1.33	0.58	0.57	0.28	0.49
		15.0	1.25	0.53	0.52	0.22	0.36
	0.5	6.0	1.84	1.17	1.16	0.99	1.81
		7.0	1.65	0.94	0.93	0.73	1.5
		10.0	1.39	0.67	0.66	0.34	0.61
		15.0	1.24	0.56	0.55	0.22	0.38
	1.0	6.0	2.43	1.91	1.88	1.31	2.45
		7.0	2.15	1.55	1.53	1.06	1.89
		10.0	1.63	0.95	0.93	0.52	1.04
		15.0	1.35	0.65	0.64	0.27	0.41
1.0	0.25	6.0	1.46	0.83	0.79	0.72	1.65
		7.0	1.33	0.72	0.68	0.55	1.27
		10.0	1.21	0.59	0.56	0.31	0.56
		15.0	1.09	0.52	0.5	0.23	0.44
	0.5	6.0	1.78	1.26	1.21	1.03	2.01
		7.0	1.57	1.02	0.98	0.75	1.46
		10.0	1.29	0.7	0.67	0.39	0.68
		15.0	1.14	0.57	0.54	0.24	0.43
	1.0	6.0	2.44	2.03	1.99	1.37	2.62
		7.0	2.1	1.66	1.62	1.1	2.15
		10.0	1.58	1.01	0.97	0.54	0.96
		15.0	1.25	0.68	0.65	0.25	0.45
2.0	0.25	6.0	1.24	1.04	0.92	0.87	1.97
		7.0	1.12	0.9	0.77	0.65	1.73
		10.0	0.93	0.69	0.57	0.39	0.73
		15.0	0.82	0.56	0.48	0.27	0.52
	0.5	6.0	1.65	1.56	1.48	1.23	2.47
		7.0	1.42	1.28	1.18	0.94	2.26
		10.0	1.08	0.87	0.75	0.45	0.98
		15.0	0.9	0.66	0.56	0.27	0.43
	1.0	6.0	2.5	2.42	2.37	1.6	2.81
		7.0	2.05	2.01	1.94	1.29	2.55
		10.0	1.41	1.28	1.18	0.62	1.35
		15.0	1.03	0.85	0.72	0.31	0.52

galaxy outflows. Moreover, the width of the $\Delta\lambda$ remains also constant ($\sim$1 Å) until $\Delta\lambda_{pix} = 0.5$ Å and $W_g = 1.0$ Å. From that point, it steadily grows. Additionally, the shape of the distribution changes with the quality of the Ly α line profile. On one hand, low values of $\Delta\lambda_{pix}$ and W_g the distribution is skewed, exhibiting a tail towards large $\Delta\lambda$ values. On the other hand, as $\Delta\lambda_{pix}$ and W_g increase, the distribution becomes more symmetric.

Next, we find that the GM-F (Fig. 11 bottom left) algorithm performance is heavily affected by the quality of the Ly α line profile. In the first place, when $\Delta\lambda_{pix}$ and W_g are low, the $\Delta\lambda$ distribution is centred around 0. This suggests that correction of the Gaussian kernel works in this range. However, as the quality of the line profile decreases, the $\Delta\lambda$ distribution moves progressively towards negative $\Delta\lambda$ values and gets broader. This is a consequence of the way in which GM-F derived the Ly α wavelength. GM-F corrects the displacement in the Ly α frequency found in GM by using a relation between the wavelength shift and the width of the red peak of the line. As the quality of the line decreases, the determination of $FWHM_{red}$ becomes more noisy and the fraction of cases with $FWHM_{red} < W_g$ rises. This causes that the $\Delta\lambda$ distributions are shifted to negative values and that it becomes broader. In fact, we find that the standard deviation of the distribution reaches 2 Å when $\Delta\lambda_{pix} = 1.0$ Å. $W_g = 2.0$ Å and $S/N_F = 6.0$.

The standard deviations of the uGM-F algorithm corresponding to redshift 3.0 and the Thin Shell geometry are listed in Table 2. uGM-F exhibits an almost identical scatter than GM-F for $W_g \leq 1.0$ Å and $\Delta\lambda_{pix} \leq 0.5$ Å. In this regime, the cases in which $FWHM_{red} < W_g$ is negligible. However, for lower qualities the $FWHM_{red} < W_g$ cases rise and increases the dispersion of GM-F. This shows that for clustering purposes, when the quality of the line profile is low, it is better to leave $FWHM_{red}$ uncorrected by the instrument point spread

Table 4. Same as Table 2 but displaying the details about the Galactic Wind geometry at redshift 2.2.

$z = 2.2$, Galactic Wind			GM	GM-F	uGM-F	NN:U	NN:B
W_g^{Rest} (Å)	$\Delta\lambda_{pix}^{Rest}$ (Å)	S/N_F	σ (Å)	σ (Å)	σ (Å)	σ (Å)	σ (Å)
0.5	0.25	6.0	1.71	0.73	0.72	0.46	0.84
		7.0	1.69	0.7	0.69	0.39	0.73
		10.0	1.65	0.65	0.65	0.3	0.53
		15.0	1.62	0.62	0.61	0.24	0.38
	0.5	6.0	1.76	0.85	0.84	0.61	1.24
		7.0	1.73	0.78	0.77	0.47	0.78
		10.0	1.63	0.68	0.68	0.35	0.57
		15.0	1.61	0.63	0.63	0.26	0.43
	1.0	6.0	1.92	1.28	1.24	0.92	1.78
		7.0	1.8	1.09	1.05	0.69	1.25
		10.0	1.67	0.85	0.82	0.4	0.65
		15.0	1.6	0.71	0.7	0.26	0.51
1.0	0.25	6.0	1.63	0.76	0.74	0.53	0.99
		7.0	1.61	0.72	0.7	0.41	0.76
		10.0	1.56	0.65	0.64	0.31	0.65
		15.0	1.53	0.61	0.6	0.23	0.41
	0.5	6.0	1.71	0.92	0.88	0.71	1.49
		7.0	1.67	0.82	0.79	0.51	1.11
		10.0	1.59	0.71	0.69	0.36	0.81
		15.0	1.53	0.64	0.63	0.27	0.47
	1.0	6.0	1.9	1.33	1.32	1.04	2.02
		7.0	1.79	1.12	1.11	0.78	1.42
		10.0	1.64	0.83	0.81	0.4	0.84
		15.0	1.55	0.7	0.68	0.3	0.52
2.0	0.25	6.0	1.52	0.92	0.83	0.71	1.91
		7.0	1.47	0.85	0.76	0.54	1.24
		10.0	1.41	0.75	0.66	0.35	0.86
		15.0	1.36	0.66	0.6	0.26	0.61
	0.5	6.0	1.63	1.17	1.13	1.01	2.3
		7.0	1.56	1.02	0.96	0.73	1.56
		10.0	1.45	0.83	0.74	0.43	1.08
		15.0	1.39	0.73	0.64	0.3	0.74
	1.0	6.0	1.92	1.79	1.76	1.36	3.12
		7.0	1.75	1.48	1.45	1.06	2.13
		10.0	1.55	1.04	0.97	0.54	1.14
		15.0	1.44	0.82	0.73	0.32	0.65

Table 5. Same as Table 2 but displaying the details about the Galactic Wind geometry at redshift 3.0.

$z = 3.0$, Galactic Wind			GM	GM-F	uGM-F	NN:U	NN:B
W_g^{Rest} (Å)	$\Delta\lambda_{pix}^{Rest}$ (Å)	S/N_F	σ (Å)	σ (Å)	σ (Å)	σ (Å)	σ (Å)
0.5	0.25	6.0	2.37	0.92	0.92	0.54	0.89
		7.0	2.32	0.86	0.87	0.47	0.74
		10.0	2.31	0.8	0.81	0.33	0.57
		15.0	2.28	0.76	0.77	0.25	0.43
	0.5	6.0	2.39	1.02	1.02	0.69	1.42
		7.0	2.34	0.94	0.94	0.53	0.99
		10.0	2.29	0.84	0.84	0.39	0.61
		15.0	2.26	0.78	0.79	0.28	0.48
	1.0	6.0	2.5	1.44	1.4	0.99	2.2
		7.0	2.43	1.24	1.21	0.75	1.49
		10.0	2.32	1.0	0.98	0.46	0.85
		15.0	2.25	0.87	0.86	0.33	0.54
1.0	0.25	6.0	2.27	0.93	0.92	0.56	0.98
		7.0	2.27	0.89	0.89	0.46	0.89
		10.0	2.22	0.8	0.81	0.36	0.66
		15.0	2.22	0.76	0.78	0.28	0.5
	0.5	6.0	2.33	1.09	1.07	0.78	1.68
		7.0	2.28	0.98	0.97	0.63	1.12
		10.0	2.22	0.85	0.85	0.38	0.7
		15.0	2.19	0.79	0.8	0.29	0.52
	1.0	6.0	2.46	1.5	1.49	1.12	2.76
		7.0	2.35	1.28	1.28	0.86	1.77
		10.0	2.26	0.99	0.99	0.5	0.7
		15.0	2.2	0.85	0.85	0.33	0.55
2.0	0.25	6.0	2.13	1.06	1.02	0.8	1.74
		7.0	2.1	0.99	0.95	0.62	1.39
		10.0	2.04	0.88	0.85	0.41	0.94
		15.0	2.0	0.79	0.8	0.31	0.6
	0.5	6.0	2.23	1.32	1.3	1.1	2.38
		7.0	2.15	1.16	1.13	0.81	1.81
		10.0	2.08	0.98	0.94	0.49	1.01
		15.0	2.0	0.85	0.83	0.32	0.73
	1.0	6.0	2.48	1.93	1.92	1.43	2.78
		7.0	2.33	1.64	1.62	1.1	2.79
		10.0	2.14	1.17	1.14	0.61	1.35
		15.0	2.05	0.96	0.93	0.37	0.71

function, although this creates a systematic bias in the mean of the $\Delta\lambda$ towards negative values.

Moreover, among the two neural network algorithms studied in this work, NN:Bright (Fig. 11 bottom right) is the one which is the most affected by the decrease of quality in the line profile. In general, the shape of the $\Delta\lambda$ distribution is composed by a prominent peak located at $\Delta\lambda = 0$ and extended wings bluewards and redwards of Lyα. Increasing $\Delta\lambda_{pix}$ and W_g causes that the wings become more elongated, and hence the accuracy of the algorithm is reduced. Meanwhile, for a set of fixed $\Delta\lambda_{pix}$ and W_g, decreasing S/N_F lowers the peak contribution while the wings remain constant, which increases significantly the width of the distributions. The large dependence on S/N_F comes from the fact that NN:Bright uses as a training set the 10 per cent brightest LAEs. As a consequence, this algorithm is trained to reproduce lines with a much higher signal-to-noise ratio than average in the LAE population. This has little effect when the faintest LAE in the sample has $S/N_F = 15$, since all the galaxy population is going to have a very good S/N and quality. In detail, when $S/N_F = 15$, the S/N of the brightest LAEs is $\sim$1000, but the difference in quality between the faint and the bright ends is very small, as the noise for the faintest galaxy is already very tiny. In other words, the quality of the Lyα line profile used for the training set is very similar to the one of the whole LAE sample, even though they exhibit a quite different S/N. However, when S/N_F decreases the differences in quality in the Lyα line profile become larger, as the faintest LAEs become more and more noisier, while the quality of bright LAEs remains almost unchanged.

Also, we find that NN:Uniform (Fig. 11 top right) is the algorithm with the best performance in most of the range of the $\{\Delta\lambda_{pix}, W_g, S/N_F\}$ volume studied here. Increasing $\Delta\lambda_{pix}$ and W_g and decreasing $S/N_F = 15$ reduce the performance of NN:Uniform. However, it is remarkable how little the spread of the $\Delta\lambda$ distribution is increased through the varied $\Delta\lambda_{pix}$ and W_g range when $S/N_F = 15$ is kept fixed. In fact $\sigma(\Delta\lambda)$ varies from 0.2 Å in the best case to only 0.35 Å in the

Table 6. Same as Table 2 but displaying the details about the Thin Shell geometry at redshift 5.7.

$z = 5.7$, Thin Shell			GM	GM-F	uGM-F	NN:U	NN:B
W_g^{Rest} (Å)	$\Delta\lambda_{pix}^{Rest}$ (Å)	S/N_F	σ (Å)	σ (Å)	σ (Å)	σ (Å)	σ (Å)
0.5	0.25	6.0	1.9	1.03	1.01	0.7	1.46
		7.0	1.9	0.98	0.97	0.59	1.13
		10.0	1.93	0.94	0.92	0.37	0.76
		15.0	2.0	0.93	0.91	0.26	0.48
	0.5	6.0	1.9	1.2	1.18	0.92	1.76
		7.0	1.86	1.09	1.08	0.7	1.3
		10.0	1.84	0.97	0.95	0.46	0.75
		15.0	1.86	0.92	0.91	0.28	0.47
	1.0	6.0	2.07	1.62	1.59	1.22	2.4
		7.0	1.95	1.41	1.38	0.92	1.83
		10.0	1.8	1.09	1.07	0.54	0.93
		15.0	1.76	0.94	0.92	0.36	0.68
1.0	0.25	6.0	1.77	1.04	0.98	0.73	1.69
		7.0	1.74	0.98	0.92	0.59	1.38
		10.0	1.75	0.92	0.86	0.41	0.8
		15.0	1.78	0.9	0.84	0.3	0.63
	0.5	6.0	1.81	1.24	1.18	0.96	1.99
		7.0	1.74	1.13	1.07	0.78	1.51
		10.0	1.71	0.96	0.9	0.46	0.82
		15.0	1.69	0.9	0.84	0.32	0.63
	1.0	6.0	2.06	1.67	1.64	1.29	2.72
		7.0	1.9	1.46	1.43	1.03	2.05
		10.0	1.71	1.1	1.06	0.55	1.0
		15.0	1.65	0.92	0.88	0.35	0.76
2.0	0.25	6.0	1.42	1.09	0.97	0.83	2.13
		7.0	1.39	1.0	0.87	0.68	1.58
		10.0	1.29	0.88	0.73	0.46	0.86
		15.0	1.21	0.82	0.67	0.36	0.73
	0.5	6.0	1.61	1.42	1.33	1.08	2.63
		7.0	1.51	1.23	1.13	0.85	1.92
		10.0	1.36	0.98	0.85	0.48	1.03
		15.0	1.24	0.86	0.71	0.34	0.7
	1.0	6.0	2.03	2.0	1.95	1.44	3.41
		7.0	1.78	1.72	1.65	1.18	2.62
		10.0	1.49	1.23	1.12	0.62	1.23
		15.0	1.31	0.95	0.82	0.37	0.76

Table 7. Same as Table 2 but displaying the details about the Galactic Wind geometry at redshift 5.7.

$z = 5.7$, Galactic Wind			GM	GM-F	uGM-F	NN:U	NN:B
W_g^{Rest} (Å)	$\Delta\lambda_{pix}^{Rest}$ (Å)	S/N_F	σ (Å)	σ (Å)	σ (Å)	σ (Å)	σ (Å)
0.5	0.25	6.0	2.92	0.87	0.86	0.54	1.06
		7.0	2.95	0.85	0.85	0.44	0.84
		10.0	3.03	0.85	0.84	0.32	0.67
		15.0	3.11	0.85	0.84	0.25	0.43
	0.5	6.0	2.77	0.93	0.92	0.65	1.26
		7.0	2.78	0.88	0.88	0.53	0.9
		10.0	2.84	0.84	0.84	0.36	0.68
		15.0	2.93	0.83	0.83	0.28	0.5
	1.0	6.0	2.71	1.19	1.16	0.86	1.52
		7.0	2.7	1.09	1.06	0.68	1.26
		10.0	2.74	0.94	0.91	0.39	0.87
		15.0	2.8	0.88	0.86	0.28	0.65
1.0	0.25	6.0	2.8	0.86	0.85	0.57	1.65
		7.0	2.81	0.84	0.83	0.49	1.01
		10.0	2.89	0.83	0.82	0.37	0.8
		15.0	2.97	0.83	0.82	0.3	0.59
	0.5	6.0	2.65	0.95	0.93	0.69	1.34
		7.0	2.66	0.89	0.87	0.55	1.12
		10.0	2.72	0.83	0.82	0.38	0.87
		15.0	2.81	0.82	0.81	0.3	0.61
	1.0	6.0	2.58	1.22	1.22	0.98	1.87
		7.0	2.58	1.07	1.08	0.72	1.4
		10.0	2.58	0.88	0.89	0.4	1.12
		15.0	2.65	0.82	0.83	0.3	0.84
2.0	0.25	6.0	2.22	0.95	0.89	0.64	2.0
		7.0	2.23	0.91	0.85	0.57	1.51
		10.0	2.25	0.85	0.8	0.38	1.06
		15.0	2.29	0.82	0.79	0.36	1.32
	0.5	6.0	2.21	1.13	1.09	0.89	2.19
		7.0	2.19	1.02	0.98	0.67	1.44
		10.0	2.19	0.89	0.84	0.4	0.99
		15.0	2.2	0.83	0.79	0.32	1.12
	1.0	6.0	2.26	1.56	1.54	1.31	2.68
		7.0	2.21	1.36	1.33	0.97	2.06
		10.0	2.14	1.02	0.98	0.51	1.07
		15.0	2.13	0.86	0.83	0.32	1.01

worst scenario. This highlights that this level of pixelization and line diluting (due to a FWHM $\neq 0$), the Ly α line profiles still contain the necessary information to identify the true Ly α wavelength. However, when S/N_F is reduced, the noise level increases and destroys part of this information. In particular, the higher $\Delta\lambda_{pix}$ and W_g values, the more likely is to lose this information due to the noise.

Finally, as shown in Tables 3, 4, 5, 6, and 7. the NN:Uniform is the best methodology at the other redshifts and geometries too.

6.3 Monopole artefacts

In general, the misidentification of the Ly α wavelength translates into an incorrect redshift determination and, hence, into a shift in the position of the LAE along the line of sight. This has a direct impact in the measured clustering of LAEs on small scales, as we showed in Section 5.2. Also, we have just shown in Section 6.2 how the quality of a given set of observed Ly α line profiles is mirrored into the identification of the Ly α wavelength. In fact, the lower the quality, the more spread the $\Delta\lambda$ distribution becomes. Here, we study how the quality of the Ly α line profiles imprints the clustering on small scales, focusing on the 2PCF. In the following, we will show the results only for two extreme cases, i.e. the worst (GM) and the best (NN:Uniform) algorithms.

In Fig. 12, we show the monopole 2PCF for all the different combinations of $\{\Delta\lambda_{pix}, W_g, S/N_F\}$ including the contribution of peculiar velocities and Ly α misidentification (through equation 12) given by the NN:Uniform algorithm. Overall, we find a clustering suppression on small scales (ξ_0, coloured lines). At larger scales, the observed LAE clustering converges to the clustering that would be observed if the Ly α frequency was known at infinite precision ($\xi_{0,R}$, dashed black line). For the GM algorithm, we find that there is a strong clustering suppression across the quality range studied here ($\gtrsim 20$ per cent at $s \sim 1$ cMpc h^{-1}). Additionally, the suppression depends slight on the quality of the Ly α line profiles. In fact,

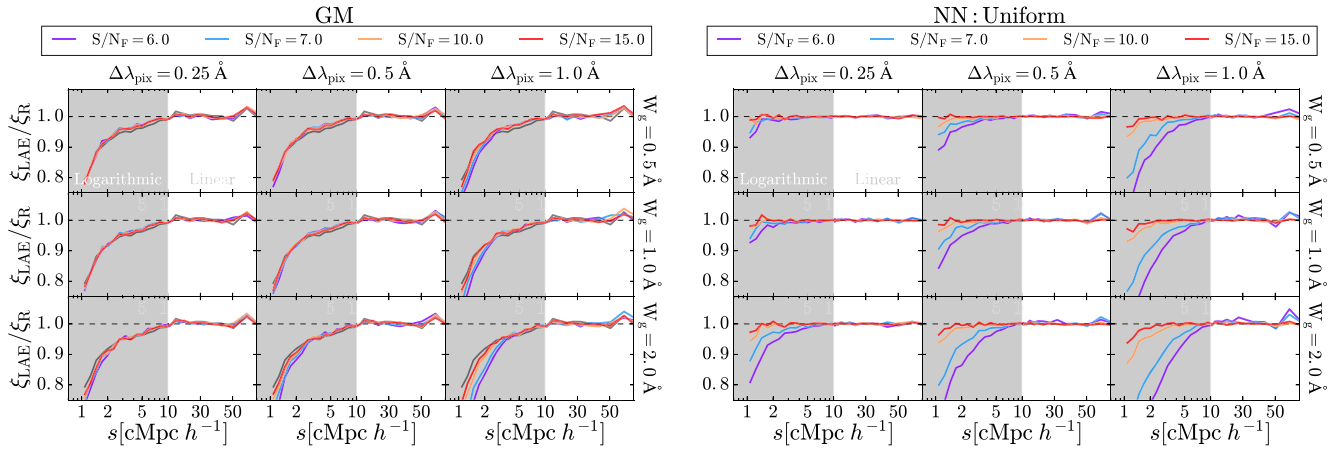

Figure 12. Ratio between the observed monopole of LAE samples with (ξ_{LAE}) and without (ξ_R) the misidentification of the Ly α frequency using the GM (left) and NN:Uniform (right) algorithms. This corresponds to the analysis of the snapshot at redshift 3.0 using the Thin Shell. The columns show $\Delta\lambda_{pix}$ = 0.25, 0.5, and 1.0 Å from left to right. Meanwhile, rows show W$_g$ = 0.5, 1.0, and 2.0 Å from top to bottom. Additionally, in each panel, S/N$_F$ = 6.0 is displayed in purple, 7.0 in blue, 10.0 in orange, and 15.0 in red. The solid grey line is the monopole computed when the ideal profiles are used (same as Fig. 8). The dashed black line signalizes unity.

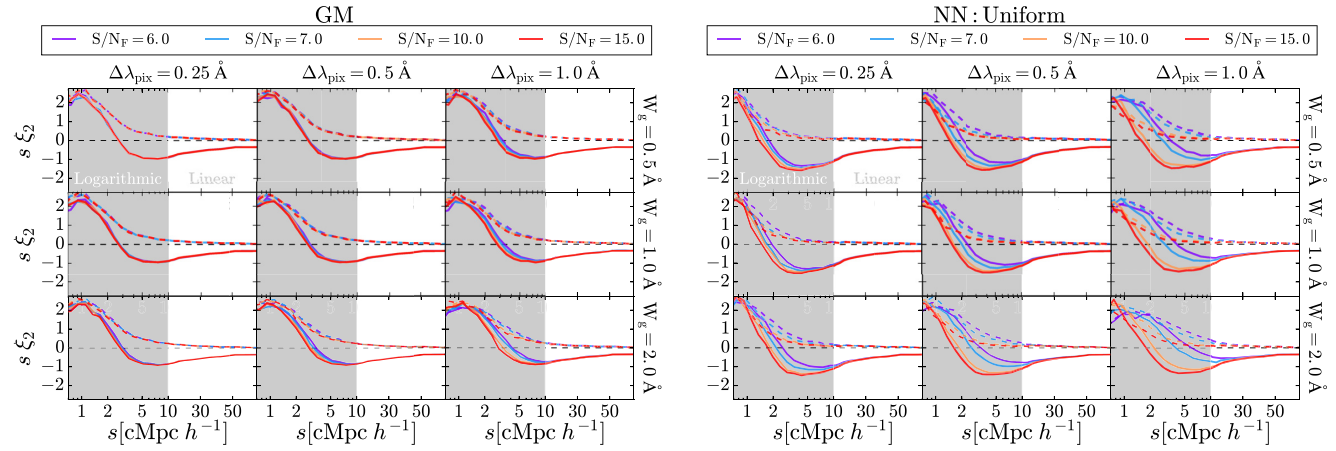

Figure 13. Quadrupole of LAE samples with (solid) and without (dashed) the peculiar motion of galaxies. The plot structure is the same as Fig. 12. Here the black dashed line indicates zero. This corresponds to the analysis of the snapshot at redshift 3.0 using the Thin Shell.

the suppression remains quite constant through all the {$\Delta\lambda_{pix}$, W$_g$, S/N$_F$} range, expect at $\Delta\lambda_{pix}$ = 1 Å and W$_g$ = 2 Å, where the lower S/N$_F$, the stronger the suppression. In fact, we also show the monopole 2PCF for ideal line profiles (solid grey curve) to illustrate how little the clustering recovered by the GM algorithms is perturbed by the line quality. This shows that the Gaussian fitting that we are applying after downgrading the quality to recover the global maximum of the line profiles works well.

For the NN:Uniform (Fig. 12 right), the convergence scale depends strongly on the quality of the observed Ly α line profiles. In general, the lower the quality, (i.e. the greater $\Delta\lambda_{pix}$ and W$_g$, and the lower S/N$_F$) the larger the clustering suppression and hence the larger the convergence scale. It is remarkable how well the NN:Uniform algorithm performs when the signal-to-noise ratio of the Ly α spectrum is good. In fact, the monopole is affected only on scales lower than 2cMpc h^{-1} when S/N$_F$ = 15. This highlights that, although diluted, the Ly α line profiles still contain the information about the Ly α wavelength. However, the addition of noise easily destroys progressively this information, causing a greater suppression. In general, the higher {$\Delta\lambda_{pix}$, W$_g$}, the most sensitive to noise becomes the clustering of LAEs.

6.4 Quadrupole artefacts

In Fig. 13, we display the quadrupole for all the multiple quality configuration for the GM and NN:Uniform algorithms, respectively. At the same time we show the samples including only the Ly α misidentification shift along the line of sight (dashed lines) and the samples including also the shift due to peculiar galaxy velocities (solid lines). In general, we find similar trends to the monopole. In particular, the lower the quality of a given set of Ly α line profiles, the larger is the clustering suppression along the line of sight (and the more positive the quadrupole becomes). However, the quadrupole appears to be more sensitive to the quality of the observed line profiles. In fact, for a given algorithm and quality configuration, similarly to the ideal case, the suppression in the quadrupole extents to larger scales than in the monopole.

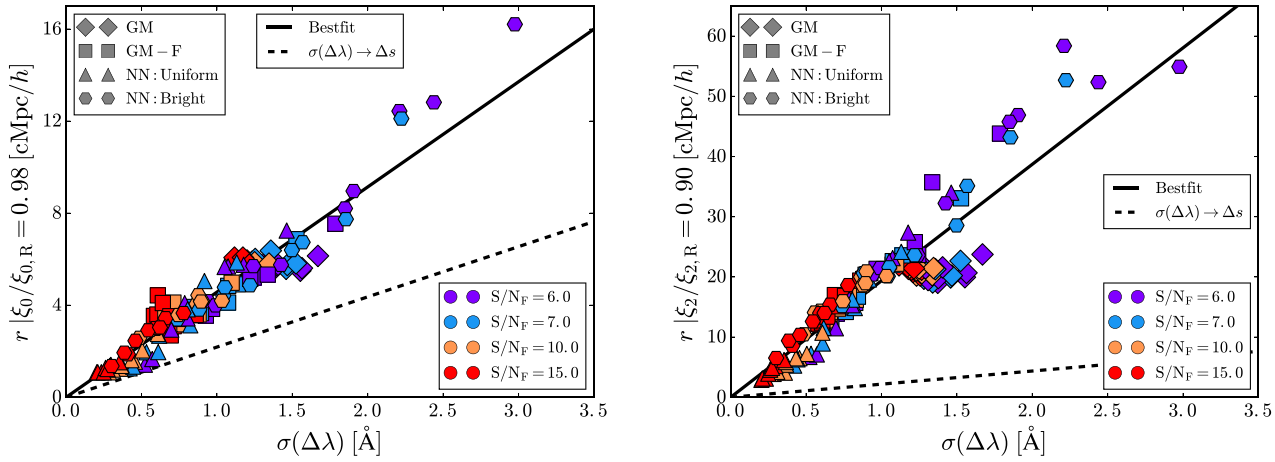

Figure 14. Left: Monopole convergence scale defined such as $\xi_0/|\mathrm{xi}_{0,R} = 0.98$ as a function of the standard deviation of the $\Delta\lambda$ distribution for all our quality configurations and Ly α identification algorithms. The GM are displayed as diamonds, GM-F as squares, NN:Uniform as triangles, and NN:Bright as hexagons. Then, the S/N_F of the sample is coloured coded. Purple means $S/N_F = 6.0$, blue 7.0, orange 10.0, and red 15.0. Additionally, we show the linear best fit to these data points in black solid line. The black dashed line indicates the direct conversion from $\Delta\lambda$ to distance shift due to Ly α misidentification. Right: Same as left, but for the quadrupole. The convergence distance for the quadrupole is defined as $\xi_2/|\mathrm{xi}_{0,R} = 0.90$.

In detail, the quadrupole of the LAE samples for GM (Fig. 13) exhibits a significant suppression of the clustering along the line of sight, which increases its amplitude. As well as in the monopole, the quadrupole of LAE samples using the GM algorithm is only slightly affected by the quality of the Ly α line profile. In fact, for most of the $\{\Delta\lambda_{\rm pix}, W_g, S/N_F\}$ combinations, the quadrupole remains almost unchanged. Although, for the worst combination of $\Delta\lambda_{\rm pix}$ and W_g studied here, the amplitude of the quadrupole correlates with S/N_F on scales smaller than $\sim 20\,{\rm cMpc}\,h^{-1}$.

Moreover, NN:Uniform (Fig. 13) exhibits the best performance over most of the $\{\Delta\lambda_{\rm pix}, W_g, S/N_F\}$ volume covered here. In contrast to the monopole, where the NN:Uniform algorithm provided a measurement without a suppression above $\sim 2\,{\rm cMpc}\,h^{-1}$ for $S/N_F = 15$, the quadrupole convergence scale for this S/N_F is extended to $\sim 5\,{\rm cMpc}\,h^{-1}$ in general. In agreement with the monopole, the clustering suppression along the line of sight in this algorithm depends strongly in the quality of the line. In this way, the suppression gets larger as the quality gets lower.

7 DISCUSSION

7.1 Relation between Ly α stack line profile and galaxy properties

As studied in Section 3, our model reproduces the trends observed between $\Delta\lambda$ and different galaxy properties. Observational studies have split the Ly α stacked line profile as function of several galaxy properties. For example, Guaita et al. (2017) studied the stacked spectrum of LAEs with spectroscopic observations from VIMOS ULTRA-DEEP SURVEY (VUDS; Le Fèvre et al. 2015; Tasca et al. 2017). Their sample consisted on 76 galaxies between redshifts $z = 2$ and $z = 4$, exhibiting both Ly α and C III] (1908 Å) as emission lines. The galaxy systemic redshifts were determined by the observed wavelength of the C III] lines. Then, authors split their galaxy sample by different observed properties. They found that $\Delta\lambda$ anticorrelated with the Ly α rest-frame equivalent width and stellar mass, while it correlated with the galaxy overdensity. These trends are in agreements with our model predictions. Guaita et al. (2017) also found that the width of the Ly α stacked line profile anticorrelated with the Ly α rest-frame equivalent width and stellar mass, as our model predicts too. Moreover, Muzahid et al. (2019) studied the relation between $\Delta\lambda$ and the SFR in 96 LAEs at $z \sim 3$ with spectroscopic observations with MUSE (Bacon et al. 2010). Muzahid et al. (2019) found the anticorrelation between $\Delta\lambda$ and the Ly α EW as well as a correlation between $\Delta\lambda$ and the SFR and Ly α luminosity.

7.2 Clustering convergence scale

In Fig. 14 (left), we show the scale where ξ_0 and $\xi_{0,R}$ converge as a function of the standard deviation of the PDF of $\Delta\lambda$ ($\sigma(\Delta\lambda)$) for all the algorithms and Ly α line profile quality configuration explored in this work. Here, we define the monopole convergence scale as the scale at which $\xi_0/\xi_{0,R} = 0.98$. Additionally, we show the best-fitting linear relation between this convergence scale and $\sigma(\Delta\lambda)$ with a slope $4.56\,{\rm cMpc}\,h^{-1}$ Å and null origin. We find that our samples follow quite well this linear relation. In detail, the samples with good S/N_F tend to cluster at lower $\sigma(\Delta\lambda)$ values and vice versa. As a reference, we also show the direct conversion from $\sigma(\Delta\lambda)$ to distance (dashed black line), computed using equation (13) and assuming $X_{\rm LoS} = 0$ and $V_{\rm LoS} = 0$. The monopole convergence distance (as defined here) has a slope a factor of $\sim$2 larger than the conversion from $\sigma(\Delta\lambda)$ to distance. Fig. 14 illustrates the strong parallelism between the behaviours of the PDF of $\Delta\lambda$ and the observed clustering on small scales (Byrohl et al. 2019).

Then, in the right-hand panel of Fig. 14 we show the relation between the convergence scale for the quadrupole (defined as $r|\xi_2/\xi_{2,R} = 0.9$). We find the same trends than in the monopole case. Basically, the lower the accuracy identifying the Ly α frequency, the larger the suppression on the observed LAE quadrupole. Additionally, the convergence scale for a fix set $\{\Delta\lambda_{\rm pix}, W_g, S/N_F\}$ is larger in the quadrupole than in the monopole, even though the definition of convergence scale is more relaxed in the quadrupole. In fact, following the same procedure as in the monopole we fitted a one degree polynomial with null origin ordinate to our samples. We find an slope of $19.2\,{\rm cMpc}\,h^{-1}$ Å, which is a factor $\sim$10 larger than the direct conversion between $\sigma(\Delta\lambda)$ and distance.

Overall, the performance in determining the Ly α wavelength is fundamentally limited by the spectral quality. A close comparison between the different models implemented in this work shows that the methodology with the lowest monopole suppression on small scales, in general, is NN:Uniform. Then it is followed by NN:Bright for high signal-to-noise ratios. However, NN:Bright is the methodology the most affected by the reduction on S/N_F. In fact, for low S/N_F NN:Bright gives the worst results (also depending on the $\Delta\lambda_{pix}$, W_g values). Then, in terms of general performance GM-F works better than GM. However, for very low quality Ly α line profiles ($\Delta\lambda_{pix} = 1.0$ Å, $W_g = 2.0$ Å, $S/N_F = 6.0$) the monopole suppression for the LAE samples identified using GM is lowest among its counterparts using different algorithms. This change of trend shows that eventually, for very low quality spectrum, finding the global maximum and setting it as the Ly α wavelength is the most robust proceeding.

7.3 Implications for HETDEX

The HETDEX (Hill et al. 2008; Adams et al. 2011) is a spectroscopic survey chasing LAEs between redshifts ∼1.9 and ∼3.5. In principle, HETDEX relies only in the measured Ly α line profile to assign a redshift to an LAE. Therefore, the LAE clustering measured by HETDEX would be sensitive to the clustering distortions studied in this work.

The typical spectral resolution and pixel size of HETDEX observations are, respectively, 5 and 2 Å in the observed frame. As the spectral quality is fixed at the observed frame, the spectral quality in the LAE's rest frame depends on the LAE redshift. Considering HETDEX's redshift range, the spectral resolution varies between ∼1.7 and ∼1.1 Å and in the LAE's rest frame. Meanwhile, the pixel size ranges from ∼0.66 to ∼0.44 Å rest frame. For these values of rest-frame spectral resolution and pixel size, we find that the NN:Uniform algorithm exhibits a better performance than the GM and GM-F algorithms. Thus, in principle, HETDEX would benefit from following the machine learning approach presented in this work.

However, it is challenging to compute the precise impact in the clustering in HETDEX. Since HETDEX LAEs populate smoothly over a given redshift window, the spectral quality is slightly different for every LAE and evolves with redshift. We plan to study this in a follow-up paper in which we will implement LAEs in a simulation light-cone. Meanwhile, here we just give a brief calculation of how much the recovered clustering can improve in HETDEX by using our methodology.

If we consider that HETDEX observations will exhibit an average $W_g = 1.5$ Å and $\Delta\lambda_{pix} = 0.5$ Å (rest frame) and $S/N_F = 6.0$. Then, the $\sigma(\Delta\lambda)$ values[4] for the GM, GM-F, NN:Uniform, and NN:Bright algorithms are 1.39, 1.02, 0.93, and 1.43 Å, respectively. For this particular configuration, the NN:Bright algorithm is outperformed by the other algorithms. This is caused by the fact that NN:Bright is only trained with the brightest LAEs. Also, GM-F performs better than GM, as in general. Meanwhile, the NN:Uniform is algorithm with the highest accuracy, exhibiting a 10 per cent better performance than GM-F. Following our results in the previous subsection, these $\sigma(\Delta\lambda)$ values translate into a convergence scale ($r|\xi_0/\xi_{0,R} = 0.98$) for the monopole of 6.33, 4.65, 4.24, and 6.52 cMpc h^{-1} for the GM, GM-F, NN:Uniform, and NN:Bright algorithms, respectively. Meanwhile, the convergence scale of the quadrupole ($r|\xi_0/\xi_{0,R} = 0.98$) are 26.7, 19.6, 17.9, and 27.5 cMpc h^{-1}.

[4]These values are calculated taking the mean of $\sigma(\Delta\lambda)$ for $W_g = 1.0$ Å and $W_g = 2.0$ Å rest frame.

7.4 Training set sizes

We acknowledge that obtaining the training sample used in Sections 5 and 6 would be extremely challenging nowadays. However, we argue that this is mainly because the LAE population presented in this series of works is, by itself, extremely challenging to obtain. Given our number density cut, each model contains 637 444 LAEs down to a Lyman α luminosity limit of $\sim 10^{42}$ erg s^{-1}, depending on redshift.

It is important to note that the training set size necessary to reach convergence depends on the LAE population for which one wants to estimate the systemic redshift. In general, a given LAE population with a larger diversity of lines would need a larger training set. For example, if you would consider LAEs in a narrow Lyman α luminosity range, most probably, the diversity of lines would be smaller than in our sample. For this kind of population, the number of galaxies with a well-constrained systemic redshift necessary to reach convergence would be smaller. In principle, this would be in general the case as Guaita et al. (2017) and Muzahid et al. (2019) found that the line profile properties depend on different galaxy properties such as the Lyman α luminosity, or the equivalent width.

Moreover, neither the NN:Uniform or the NN:Bright training samples are selected optimally. In Sections 4, 5, and 6, it is made clear that the NN:Uniform performs much better than NN:Bright due to the difference in training sets. However, there is still plenty of space for improvement in the way of selecting the training sample. For example, a more efficient selection criteria would be to perform a latin hypercube sampling of different properties, such as the Lyman α luminosity, the FWHM of the line, or its global maximum. In principle, this could lead to an faster convergence, in terms of $N_{LAE}^{Training}$, in comparison with NN:U.

Also, in general, it is not necessary to reach convergence to improve the results from previous methodologies. For example, focusing in the ideal case using the Thin Shell at redshift 3.0 (Fig. 4), the performance of the best traditional methodology (in this case GM-F) gives $\Delta\lambda = 0.52$ Å (see Table 1). The performance of NN:U in the range of training set sizes considered (down to 100 galaxies) is, in almost every case, better than $\Delta\lambda = 0.52$ Å. This is also the case for the Galactic Wind.

Finally, we explore the Ly α wavelength determination accuracy as a function of the training set size and line profile quality in Fig. 15 for the Thin Shell at redshift 3.0 and using the NN:Uniform algorithm. For this, we picked two of the $\{W_g, \Delta\lambda_{pix}, S/N_F\}$ combinations studied in Section 6: {1.0, 0.5, 7.0 Å} (blue) and {0.5, 0.25, 7.0 Å} (red). We also added {0.25, 0.125, 7.0 Å} (brown) and {0.1, 0.05, 7.0 Å} (purple) in order to have a smooth transition from realistic to ideal line profiles. We kept $S/N_F = 7.0$ fixed because it is a low intermediate value. We have checked that we obtain qualitatively the same trends with the other S/N_F values. Overall, we find that as we increase the line profile quality, $\sigma(\Delta\lambda)$ decreases for fix $N_{LAE}^{Training}$ values. This means that the better line profile quality, the smaller the training sample has to be in order to obtain the same accuracy.

For comparison, we show $\sigma(\Delta\lambda)$ of the best standard methodology for these configurations, GM-F, in horizontal dashed lines. For the training set size used in Section 5 and Section 6 (vertical black dashed line), NN:Uniform always exhibits a better accuracy than GM-F, as shown previously. However, the accuracy of NN:Uniform depends on the training set sample. If $N_{LAE}^{Training}$ is too small, the accuracy of GM-F becomes better than the one of NN:Uniform. The $N_{LAE}^{Training}$

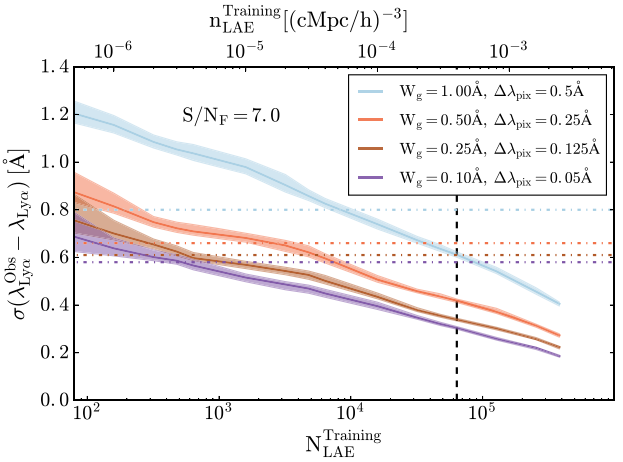

Figure 15. Accuracy of NN:Uniform at redshift 3.0 using the Thin Shell geometry. Each colour indicates a line profile quality configuration $\{W_g, \Delta\lambda_{pix}\}$ keeping $S/N_F = 7.0$, being blue {1.0, 0.5 Å}, red {0.5, 0.25 Å}, brown {0.25, 0.125 Å}, and purple {0.1, 0.05 Å}. For each value of $n_{LAE}^{Training}$, we compute 100 iterations if $n_{LAE}^{Training} \leq 4 \times 10^{-4}(cMpc/h)^{-3}$ and 20 if $n_{LAE}^{Training} > 4 \times 10^{-4}(cMpc/h)^{-3}$. The solid lines show the median and the shaded regions of the 1σ scatter. The horizontal coloured dashed lines show the accuracy of GM-F for each quality. The vertical black line shows the fiducial value of $n_{LAE}^{Training}$.

range in which NN:Uniform performs better than GM-F depends on the line profile quality. In fact, the better the line profile quality, the lower $N_{LAE}^{Training}$ is necessary to improve GM-F. For example, for the case with $\{W_g, \Delta\lambda_{pix}, S/N_F\} = \{0.1, 0.05, 7.0$ Å$\}$, about 500 LAEs with a good redshift determination are required to improve GM-F. However, for a sample with $\{W_g, \Delta\lambda_{pix}, S/N_F\} = \{1.0, 0.5, 7.0$ Å$\}$, around 10^4 LAEs would be necessary.

8 CONCLUSIONS

In this work, we have addressed the clustering distortions in LAE samples due to the misidentification of the Ly α wavelength in the Ly α line profile and how to mitigate them using neural networks. With this goal, we have analysed the Ly α line profiles from our previous LAE theoretical model (Gurung-López et al. 2020) that includes the Ly α RT in the ISM and in the IGM. Our LAE model reproduces by construction the observed LAE luminosity function and a bunch of their observables, such us the Ly α escape fraction, the metallicity distribution, and clustering amplitude of LAEs (Gurung-López et al. 2019a).

After analysing the stacked Ly α line profiles of our LAE model we find that:

(i) The stacked Ly α line profile is affected by the IGM. In particular, as the IGM Ly α optical depth increases with redshift, the higher the redshift, the more the stacked Ly α line profile is modified by the IGM. At low redshifts ($z = 2.2$ and $z = 3.0$), we find that the stacked Ly α line profile changes slightly and only bluewards the Ly α wavelength. Meanwhile, at redshift 5.7, the radiative transfer in the IGM modifies the stacked Ly α line profile up to 2 Å redder than the Ly α wavelength.

(ii) The stacked Ly α line profile depends on both galaxy and IGM properties. On one hand, our model predicts that for the Thin Shell and the Galactic Wind outflow geometries, the Ly α stacked line profile is centred at large wavelength for higher values of SFR and $L_{Ly\alpha}$. Additionally, LAEs with higher values of Ly α equivalent width exhibit a bluer Ly α stacked line profile. Meanwhile, in the Galactic Wind we find only a small dependence on the stellar mass, while in the Thin Shell the lower the M_*, the more redshifted is the peak of the stacked Ly α line profile. Over all, these trends are in good agreement with observational works (Guaita et al. 2017; Muzahid et al. 2019). On the other hand, the stacked Ly α line profiles are more redshifted in high-density environments and low in IGM large-scale line-of-sight velocity, its gradient, and density gradient.

Then, we have introduced a novel approach to measure the systemic redshift of LAEs from their Ly α line using neural networks. In this frame, given a survey that only observed the Ly α line, a fraction of the LAE population could be reobserved to find other features to determine their systemic redshifts. Then, this subpopulation would be used to train a neural network that predicts the systemic redshift. In particular, we have explored two different ways of building the training set: (i) the reobserved sources are chosen uniformly across their properties (NN:Uniform) and (ii) only the brightest LAEs are reobserved (NN:Bright). In order to assess the performance of these methodologies, we compare them with others found in the literature. In particular, we use (i) GM that uses the global maximum of the Ly α line profile to assign a systemic redshift (Verhamme et al. 2018; Byrohl et al. 2019; Muzahid et al. 2019) and (ii) GM-F, suggested by Verhamme et al. (2018), that uses the width of the Ly α line correct for the redshift due to the Ly α RT in the ISM.

First, we focus on the Ly α line profiles produced by our model, which are ideal in terms of signal to noise and pixelization. We find that the NN:Uniform and NN:Bright algorithms perform better than GM and GM-F. In fact, the distributions of the displacement of Ly α (Δλ) is the broadest for the GM, followed by the GM-F and NN:Bright, while for the NN:Uniform, it is the thinnest. Then, we study how each of these methods impact the observed clustering of LAEs with ideal line profiles. We find that:

(i) In general, the power spectrum exhibits a damping at small scales that disappears at large enough distances, as found in Byrohl et al. (2019). This damping is directly linked to the performance recovering the systemic redshift of the LAEs. In fact, the clustering of samples using GM and GM-F exhibits a decrease of 80 per cent of power at $k_{\shortparallel} = 1.0$. In contrast, the samples using NN:Uniform and NN:Bright exhibit a much shallower damping. Typically, samples using NN:Bright have a damping of 10 per cent or less at $k_{\shortparallel} = 1.0$. Meanwhile, the samples using NN:Uniform are mostly unaffected at $k_{\shortparallel} = 1.0$, as the power spectrum damping is of the order of the 1 per cent.

(ii) The monopole also exhibits a damping at small scales parallel to the one observed in the power spectrum. In particular, the power suppression in the monopole is of the order of 1 per cent at 1 cMpc h^{-1}. Meanwhile, for GM and GM-F the monopole damping can go up to the 60 per cent at 1 cMpc h^{-1}. Also, for the GM and GM-F, the suppression extends up to 10 cMpc h^{-1}, where, in general, the intrinsic and the observed monopole converge.

(iii) The quadrupole is also sensitive to the systemic redshift determination. In fact, we find that the lower the accuracy recovering the Ly α wavelength, the more power exhibits the quadrupole between 1 and 10 cMpc h^{-1}. The quadrupole of the samples using NN:Bright and, specially, NN:Uniform are mostly identical to the quadrupole of the underling LAE population.

Next, we explore the benefits of using NN:Uniform and NN:Bright in comparison with GM and GM-F in realistic line profiles. With this goal, we lower the quality of the Ly α line profile mocking

several artefacts in observations. In practice, (i) we dilute the line assuming different instrumental FWHM, (ii) we reduce the wavelength resolution by pixelizing the line, and (iii) we include noise in the Ly α line profile. Then, we study the properties of these samples, finding:

(i) The performance, i.e. the distributions $\Delta\lambda$ is tightly connected to the spectral quality. Overall, we find that in the spectral quality range covered in this work, the NN:Uniform is the best methodology. Additionally, the NN:Bright is very affected by the noise in the spectrum. This is a result of using only the brightest LAEs as training sample, as this one lacks faint LAEs in which their Ly α line profile is noisy. In this way, NN:Bright is the second best algorithm for high signal-to-noise lines. Meanwhile, GM-F progressively gets worse as the spectral quality decreases. GM-F is mostly affected by the instrumental FWHM, as this modifies the width of the Ly α line and overcorrects the RT in the ISM. Finally, GM is the algorithm with the lowest performance in most of the spectral quality regime studied here. However, GM is quite insensitive to lowering the spectral quality, specially, increasing the noise. This, in the worst spectral quality considered here, translates into a better accuracy of GM in comparison with NN:Bright and GM-F, while it is similar to NN:Uniform.

(ii) Consequently, the monopole and the quadrupole are affected by the reduction of the spectral quality. We find that NN:Uniform is algorithm that recovers better the clustering of the underlying LAE population. However, as the spectral quality is reduced, the damping of power at small scales increases, as in GM-F and NN:Bright. Meanwhile, the clustering damping at small scales is quite constant for GM through the dynamical range of the spectral quality studied here.

(iii) There is a linear relation between the algorithm performance and the typical scale up to which the clustering power is decreased. This will be useful for the design of future surveys based on the Ly α line.

Therefore, we conclude that spectroscopic Ly α based surveys such as HETDEX might benefit from measuring the systemic redshift of a relative small subsample of LAEs, using other spectral features. And then, using this subsample to train machine learning algorithms to predict the systemic redshift of the rest of the observed LAE population.

ACKNOWLEDGEMENTS

We acknowledge and thank the great labour of the scientific referee. Their comments improved greatly the quality, clearance, and content of this work. This work has made used of CEFCA's Scientific High Performance Computing system that has been funded by the Governments of Spain and Aragón through the Fondo de Inversiones de Teruel, and the Spanish Ministry of Economy and Competitivenes (MINECO-FEDER, grant AYA2012-30789). The authors acknowledge the support of the Spanish Ministerio de Economia y Competividad project No. AYA2015-66211-C2-P-2. We acknowledge also STFC Consolidated Grants ST/L00075X/1 and ST/P000451/1 at Durham University. This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/P002293/1, ST/R002371/1, and ST/S002502/1, Durham University, and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure. SS was supported in part by JSPS KAKENHI Grant Number JP15H05896, and by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan. SS was also supported in part by the Munich Institute for Astro- and Particle Physics (MIAPP) which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC-2094-390783311).

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Adams J. J. et al., 2011, ApJS, 192, 5
Ahn S., 2003, J. Korean Astron. Soc., 36, 145
Ahn S.-H., Lee H.-W., Lee H. M., 2000, J. Korean Astron. Soc., 33, 29
Bacon R. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 7735, Ground-Based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 773508
Baugh C. M. et al., 2019, MNRAS, 483, 4922
Behrens C., Byrohl C., Saito S., Niemeyer J. C., 2018, A&A, 614, 11
Benitez N. et al., 2014, preprint (arXiv:1403.5237)
Beutler F. et al., 2014, MNRAS, 443, 1065
Bonoli S. et al., 2020, preprint (arXiv:2007.01910)
Byrohl C., Saito S., Behrens C., 2019, MNRAS, 489, 3472
Caruana J. et al., 2018, MNRAS, 473, 30
Chonis T. S. et al., 2013, ApJ, 775, 99
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Erb D. K. et al., 2014, ApJ, 795, 33
Gronke M., Dijkstra M., McCourt M., Oh S. P., 2016, ApJ, 833, L26
Gronwall C. et al., 2007, ApJ, 667, 79
Guaita L. et al., 2017, A&A, 606, A19
Gurung López S., Orsi Á. A., Bonoli S., Baugh C. M., Lacey C. G., 2019a, MNRAS, 486, 1882
Gurung-Lopez S., Orsi A. A., Bonoli S., 2019b, MNRAS, 490, 733
Gurung-López S., Orsi Á. A., Bonoli S., Baugh C. M., Lacey C. G., 2019a, MNRAS, 486, 1882
Gurung-López S., Orsi Á. A., Bonoli S., 2019b, MNRAS, 490, 733
Gurung-López S., Orsi Á. A., Bonoli S., Padilla N., Lacey C. G., Baugh C. M., 2020, MNRAS, 491, 3266
Harrington J. P., 1973, MNRAS, 162, 43
Henry A., Berg D. A., Scarlata C., Verhamme A., Erb D., 2018, ApJ, 855, 96
Hill G. J. et al., 2008, in Kodama T., Yamada T., Aoki K., eds, ASP Conf. Ser. Vol. 399, Panoramic Views of Galaxy Formation and Evolution. Astron. Soc. Pac., San Francisco, p. 115
Hu E. M., Cowie L. L., McMahon R. G., 1998, ApJ, 502, L99+
Jones T., Stark D. P., Ellis R. S., 2012, ApJ, 751, 51
Kaiser N., 1987, MNRAS, 227, 1
Konno A. et al., 2018, PASJ, 70, S16
Lacey C. G. et al., 2016, MNRAS, 462, 3854
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Laursen P., Sommer-Larsen J., Razoumov A. O., 2011, ApJ, 728, 52
Le Fèvre O. et al., 2015, A&A, 576, A79
Leclercq F. et al., 2017, A&A, 608, A8
Malhotra S., Rhoads J. E., 2002, ApJ, 565, L71
Matthee J., Sobral D., Best P., Smail I., Bian F., Darvish B., Röttgering H., Fan X., 2017, MNRAS, 471, 629
Muzahid S. et al., 2019, MNRAS, 496, 1013
Neufeld D. A., 1990, ApJ, 350, 216
Orlitová I., Verhamme A., Henry A., Scarlata C., Jaskot A., Oey M. S., Schaerer D., 2018, A&A, 616, A60
Orsi A., Lacey C. G., Baugh C. M., 2012, MNRAS, 425, 87
Ouchi M. et al., 2008, ApJS, 176, 301
Ouchi M. et al., 2018, PASJ, 70, S13
Oyarzún G. A., Blanc G. A., González V., Mateo M., Bailey III J. I., 2017, ApJ, 843, 133

Planck Collaboration XIII, 2016, A&A, 594, A13
Rhoads J. E., Malhotra S., Dey A., Stern D., Spinrad H., Jannuzi B. T., 2000, ApJ, 545, L85
Rudie G. C., Steidel C. C., Pettini M., 2012, ApJ, 757, L30
Shibuya T. et al., 2018, PASJ, 70, S15
Shoji M., Jeong D., Komatsu E., 2009, ApJ, 693, 1404
Sobral D., Matthee J., Darvish B., Schaerer D., Mobasher B., Röttgering H. J. A., Santos S., Hemmati S., 2015, ApJ, 808, 139
Sobral D. et al., 2018, MNRAS, 477, 2817
Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, ApJ, 462, L17+
Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, ApJ, 717, 289
Steidel C. C., Bogosavljević M., Shapley A. E., Kollmeier J. A., Reddy N. A., Erb D. K., Pettini M., 2011, ApJ, 736, 160
Tasca L. A. M. et al., 2017, A&A, 600, A110
Trainor R. F., Steidel C. C., Strom A. L., Rudie G. C., 2015, ApJ, 809, 89
Verhamme A., Schaerer D., Maselli A., 2006, A&A, 460, 397
Verhamme A. et al., 2018, MNRAS, 478, L60
Wyithe J. S. B., Dijkstra M., 2011, MNRAS, 415, 3929
Zheng Z., Miralda-Escudé J., 2002, ApJ, 578, 33
Zheng Z., Cen R., Trac H., Miralda-Escudé J., 2010, ApJ, 716, 574
Zheng Z., Cen R., Trac H., Miralda-Escudé J., 2011, ApJ, 726, 38

APPENDIX A: WAVELENGTH IMPORTANCE

In this section, we quantify which wavelength of the line profile contributes the most to the determination of the $\lambda_{\mathrm{Ly}\alpha}^{\mathrm{Obs}}$ in the neural networks in this work. With this goal in mind, for a given wavelength bin centred at the wavelength λ_{pix} we define its importance as

$$\mathrm{I}(\lambda_{\mathrm{pix}}) = \sigma(\lambda_{\mathrm{pix}})\ \omega(\lambda_{\mathrm{pix}}), \tag{A1}$$

where $\omega(\lambda_{\mathrm{pix}})$ is each weight for the pixel centred at λ_{pix} and $\sigma(\lambda_{\mathrm{pix}})$ is the standard deviation of that pixel across the training set.

In Fig. A1, we show the importance as a function of wavelength for the NN:Uniform at different redshifts and for the two outflow geometries. In general, we find that the wavelength range with a significant importance is wider than the Ly α stacked line profiles at $z = 2.2$ and 3.0, while at $z = 5.7$ both are similar. On one hand, at $z = 5.7$ the IGM absorbs mostly all the flux bluewards Ly α (Zheng et al. 2010; Laursen et al. 2011; Gurung-López et al. 2020). Therefore, $\lambda_{\mathrm{Ly}\alpha}$ is extracted mainly from the information closer than 2 Å to $\lambda_{\mathrm{Ly}\alpha}$. Meanwhile, the regions with $\lambda < \lambda_{\mathrm{Ly}\alpha}$ and $\lambda > \lambda_{\mathrm{Ly}\alpha} + 5$ Å contain little information in comparison. On the other hand, at lower redshifts ($z = 2.2$ and 3.0), the IGM absorption is not that strong and the neural networks need the information from a broader wavelength range. In fact, at these redshifts, most of the importance is located between $\lambda_{\mathrm{Ly}\alpha}$ and 5 Å redwards. Additionally, the spectral region with a significant importance spawns from ∼1 Å bluewards $\lambda_{\mathrm{Ly}\alpha}$ to ∼15 Å redwards $\lambda_{\mathrm{Ly}\alpha}$. This broad wavelength range exhibits the large variety of line profiles in our LAE population, as the presence of very broad lines (FWHM ∼ 10 Å) extend the importance range up to ∼15 Å.

APPENDIX B: COMPARISON OF THE ACCURACY BETWEEN MODELS

In this section, we speculate about why a given methodology works better for a certain model (combination of redshift and outflow geometry) than for other. Sometimes this can be intuitive. For example, GM-F sometimes fails at $z = 5.7$, as discussed above, due to the large IGM absorption. However, there are cases in which it is not trivial to find an answer. For example, in the ideal case at redshift 3.0 (Fig. 4), in which NN:Uniform reaches a better performance for the Thin Shell than for the Galactic Wind for large enough training sets.

We think that the cause of the different performances is the 'diversity' of lines of each model, i.e. the amount of different line profiles. For example, if a population exhibits lines with one, two, and three peaks, most likely, it would have a larger line diversity than another population containing only line with one peak. Other factor that matter for this would be the width of the distributions of the global maximum or of the FWHM of the line, among others.

In principle, the larger the variety of line profiles, the larger the dispersion of NN:Uniform might be for a fixed training set size. As explained above, our models are build so that they reproduce the LAE LF, which makes that every model has a unique distribution outflow expansion velocity V_{exp}, neutral hydrogen column density N_{H}, dust optical depth τ_a, and IGM transmission curves $T(\lambda)$.

It is challenging to quantitatively characterize the diversity of lines. For example, given a fix outflow geometry, one could think that the

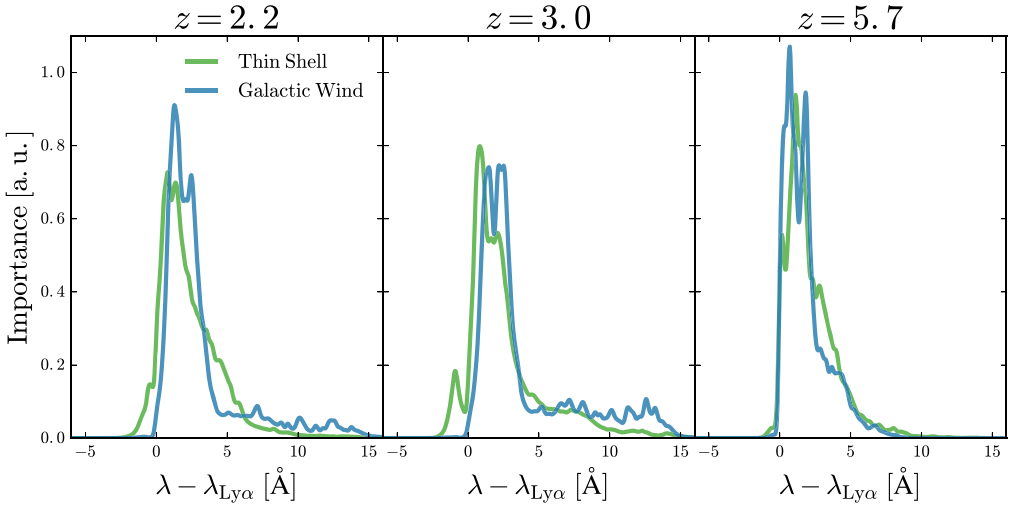

Figure A1. Importance of each wavelength bin used in the NN:Uniform algorithm at redshifts 2.2, 3.0, and 5.7 from left to right. The models using the Thin Shell outflow geometry are displayed in green, while their Galactic Wind counterparts are shown in blue.

bigger the covered volume in the 4D space (V_{exp}, N_H, τ_a, and $T(\lambda)$), the larger would be the diversity of lines, thus, the larger would be $\sigma(\Delta\lambda)$ for a fix methodology. However, this is not the case: in Gurung-López et al. (2020) (fig. A1), we showed the $\{V_{exp}, N_H, \tau_a\}$ region covered by the Thin Shell and the Galactic Wind at $z = 2.2, 3.0$ and 5.7. For both geometries, the volume occupied at $z = 5.7$ is larger than at $z = 2.2$ and $z = 3.0$. This would explain that the NN:Uniform performance in the thin shell is a 25 per cent lower at $z = 5.7$ than at the other redshifts. However, the NN:Uniform in the Galactic Wind has greater accuracy at $z = 5.7$ than at $z = 2.2$ and $z = 3.0$ in the ideal case, which does not support the above, apparently, naive reasoning. This can be explained because there are degeneracies between the different parameters. For example, between V_{exp} and N_H.

Then, the comparison between outflow geometries is even more complicated. Even if two models cover exactly the $\{V_{exp}, N_H, \tau_a, T(\lambda)\}$ region, the diversity of line profiles might be different, as most probably, the mapping between the outflow properties and line profile would be different.

Also, it is not always the case in which the accuracy of methodology is better in the Thin Shell. For example, NN:Uniform at $z = 5.7$ in the ideal case is more accurate in the Galactic Wind. At the same time, GM-F at $z = 2.2$ is more accurate for the galactic wind (as opposite to NN:Uniform at $z = 2.2$).

This paper has been typeset from a TEX/LATEX file prepared by the author.